MECHANISMS OF GENOME STABILITY AND DNA REPAIR IN *PLASMODIUM FALCIPARUM*

A Dissertation

Presented to the Faculty of the Weill Cornell Graduate School

of Medical Sciences

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Susannah Fabri Calhoun

December 2017

MECHANISMS OF GENOME STABILITY AND DNA REPAIR IN *PLASMODIUM FALCIPARUM*

Susannah Fabri Calhoun

Cornell University 2018

The human malaria parasite *Plasmodium falciparum* replicates within circulating red blood cells where it is subjected to conditions that frequently cause DNA damage. The repair of DNA double stranded breaks (DSBs) is thought to rely almost exclusively on homologous recombination (HR) due to a lack of efficient non-homologous end joining (NHEJ), however given that the parasite is haploid during this stage of its lifecycle, the mechanisms involved in maintaining genome stability are poorly understood. This work investigates methods of DNA repair and how these pathways function in maintaining the *P. falciparum* genome. First, we investigate the essentiality of Rad51 in response to DNA damage and describe how the parasite developmental stage dictates DNA repair ability based on the possibility of HR repair. Life cycle stage is shown to mediate resistance to DNA damage caused by X-ray irradiation likely due to the parasite's ability to perform HR at different stages of the erythrocytic cycle. By creating a Rad51 dominant negative mutant we find that Rad51 may not be an essential protein for HR repair and the loss of function differentially affects stage-specific repair. Secondly, we show that parasites utilize a competitive balance between de novo telomere addition, also called "telomere healing", and HR to stabilize chromosome ends. Products of both repair pathways were observed in

response to DSBs that occurred spontaneously during routine *in vitro* culture or resulting from experimentally induced DSBs, demonstrating that both pathways are active in repairing DSBs within subtelomeric regions and the pathway choice is sequence dependent. Finally, we discuss how this research fits into the field of knowledge about DNA repair in *P. falciparum* and how evolution of the DNA repair pathways has mediated plasticity of the genome.

BIOGRAPHICAL SKETCH

Susannah Fabri Calhoun grew up in Brooklyn, NY with her family. She attended NYC public schools through high school. At Midwood High School, Susannah was in the Humanities Honors Program and competed in the Intel Science Talent Search. Susannah attended Smith College in Northampton, MA where she majored in Biological Sciences and minored in African-American Studies. She discovered an interest in parasitic diseases during her study abroad at the University of Edinburgh in Scotland. During her undergraduate summers Susannah interned at both NASA Langley Research Center and Merck & Co. Post college she worked briefly at Regeneron Pharmaceuticals before becoming a Research Technician in the lab of Alessio Accardi at Weill Cornell Medical College. She began the Immunology & Microbial Pathogenesis program at Weill Cornell Graduate School of Medical Sciences in September 2012. Susannah joined the lab of Kirk Deitsch in June 2013 and completed her thesis research studying DNA repair in *Plasmodium falciparum.* She defended her thesis in October 2017, receiving her PhD in Immunology and Microbial Pathogenesis.

My thesis is dedicated to Grandear; Mary Kane, my maternal grandmother who had a love of nature and science and would be proud to see what I have achieved as a woman in science. I wish she was here to share in this accomplishment.

ACKNOWLEDGEMENTS

I first would like to thank my mentors Dr. Kirk Deitsch and Dr. Laura Kirkman. Kirk welcomed me into his lab, guided me through my thesis research, and shared with me his passion and expertise for malaria research. He was always available for advice and troubleshooting and I am forever grateful for his mentorship. Laura originated some of my projects and provided invaluable insight to their development. She has been both an excellent mentor and role model, both supporting and challenging me. Jake Reed was my partner in the *Plasmodium* PacBio work and our paper would not have been possible without his genome analysis, computer expertise and hard work. The entire Deitsch lab has been integral to both my success and happiness in the lab. I especially want to thank: Alexis Dziedziech; who created the Rad51 parasites, and Amanda Chan; who helped with the irradiation sensitivity assays. The postdocs I have worked with: Uchechi Ukaegbu, Johan Ankarklev, and Xu Zhang have provided me with science knowledge, career advice and friendship. The Mason Lab collaboration was a tremendous addition to my research and the generosity of Christopher Mason and the expertise and assistance of Noah Alexander made this project seamless. The Kafsack and Templeton labs provided valuable research help and support. My thesis committee: Dirk Schnappinger and Michael Glickman have been instrumental in the development of my research and scientific career. The Immunology and Microbial Pathogensis leadership supported me throughout my time at Weill Cornell. All of my graduate school friends have been a source of refuge, encouragement and made my PhD a wonderful experience. My friends from Brooklyn, Smith, and post-college have always provided a joyous diversion and supported me even when they didn't understand my work. And most importantly, I have eternal gratitude and love for my family: my mother, sister and uncle. My mom has always been my biggest supporter and confidant, providing the foundation on which I have built my life. My sister Annabel has been the most devoted champion of my success and my Uncle John's investment in my life has always boosted me. And in a different, but equally significant way I am grateful for the National Institute of Health AI 52390 Diversity supplement that has funded my research for 3 years.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Malaria: A Global Health burden

Malaria is an infectious disease that is endemic to tropical and subtropical regions of the world, currently 91 countries, and is transmitted by the *Anopheles* mosquito.[1, 2] This means that nearly half the world's population is at risk of malaria. In 2015, there were an estimated 212 million new cases of malaria and 429 thousand deaths.[1] Africa accounts for about 90% of malaria cases, with many children being especially susceptible to malaria illness and death. Malaria is a huge killer of children under the age of five, claiming the life of 1 child every 2 minutes.[1] Programs designed to control and ultimately eliminate malaria are being implemented and improvements in disease burden are happening worldwide. This is achieved with programs such as vector control with insecticide-treated mosquito nets and insecticide spraying, chemoprevention for pregnant women, diagnostic testing and malaria surveillance systems.[1] Gains in malaria control are jeopardized by insecticide resistance by mosquitoes and drug resistance to antimalarials, particularly to artemisinin, the most effective antimalarial available.[1, 2] There is currently no effective vaccine for malaria. Therefore, it is imperative that we comprehend the molecular biology of *Plasmodium*, the protozoan parasite that causes the disease, to determine the best target interventions to reduce morbidity and mortality attributable to malaria. The *Plasmodium* genus has five species that can infect humans: *P.falciparum, P.ovale, P. malariae, P. vivax,* and *P. knowlesi*. *P. falciparum* causes the deadliest form of malaria and is responsible for the majority of malaria cases.[1, 3]

1.2 The Life Cycle of *Plasmodium*

The lifecycle of all *Plasmodium* species is comprised of asexual replication in the liver and erythrocytes that circulate through the vertebrate host and sexual reproduction in the midgut of the *Anopheline* mosquito vector. The cycle begins when a female *Anopheles* mosquito injects sporozoites into the vertebrate host during a bloodmeal. The sporozoites enter the bloodstream and travel to the liver where they invade hepatocytes and there sustain multiple rounds of asexual replication to form thousands of merozoites. The hepatocytes rupture releasing merozoites into the bloodstream to invade erythrocytes. In the erythrocytic stage, the parasite is encased in a parasitophorous vacuole where the merozoite matures through distinct morphological stages, from ring to trophozoite to schizont. During schizogony the nucleus divides mitotically and asynchronously, then the cell segments to produce anywhere from 16-30 daughter merozoites. The erythrocyte then ruptures releasing the merozoites which re-infect new erythrocytes. This occurs in a cyclical pattern over a period of 48 hours for *P. falciparum*.[4, 5] The rupture of the schizont is associated with fever, nausea, headaches and other symptoms of a systemic pro-inflammatory cytokine response, much of which is caused by the cells of the innate immune system of the host.[4] Other severe pathologies of the disease include anemia, vascular obstruction, cerebral malaria and possibly death.[5] During the asexual cycle, a small population of parasites commit to sexual stage development and differentiate into gametocytes. The female and male gametocytes are taken up by the mosquito vector and fertilization occurs in the midgut, followed by meiosis to produce the zygote. Besides the zygote the *Plasmodium* species are haploid throughout the entire life cycle. The zygote

differentiates to a motile ookinete, breaks through the midgut wall and becomes an oocyst. Replication occurs within the oocyst forming sporozoites, which egress and travel to the salivary glands. These sporozoites will infect new hosts, beginning the cycle of infection and replication over again.[4, 5] (Figure 1.1) The genetic, structural and metabolic changes that happen to the parasite during this complex life cycle are not fully elucidated. The ability to culture both the asexual and sexual erythrocytic stages of *Plasmodium falciparum* in the laboratory, using human red blood cells and media, allows for in depth research to gain better knowledge of the parasite's biology and implementation of interventions has for malaria control and eradication.

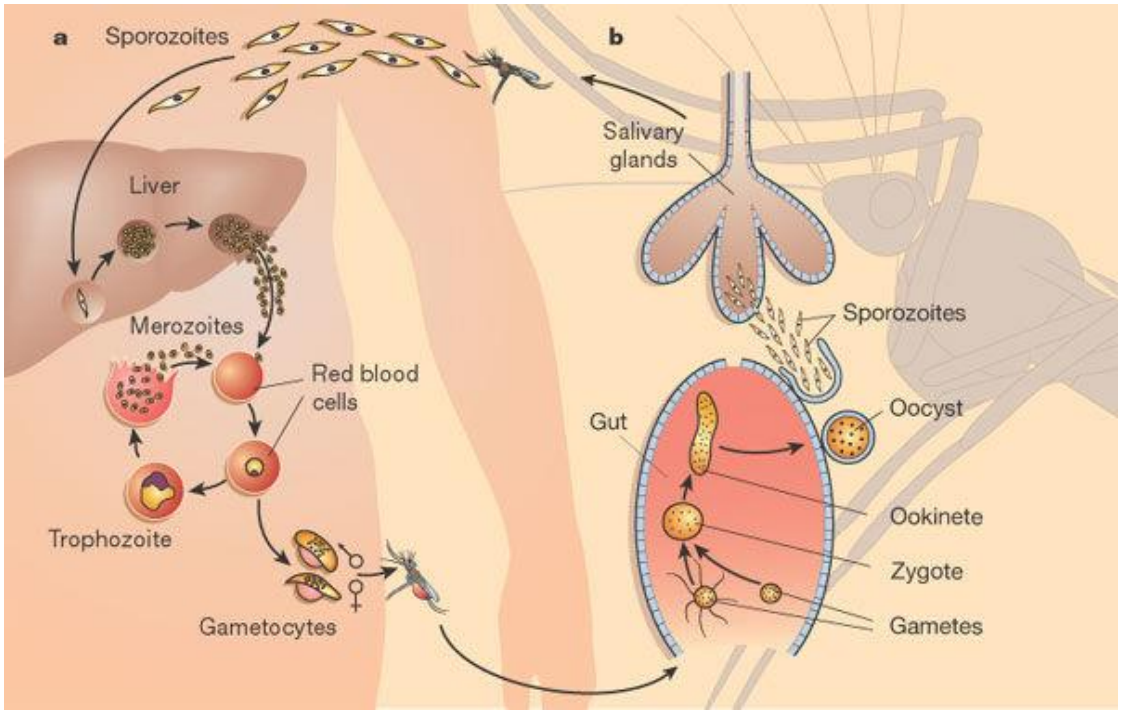


Figure 1.1. The life cycle of *Plasmodium falciparum.* A. When a parasite-infected mosquito feeds on a human, it injects the parasites in their sporozoite form. These travel to the liver, where they develop through several stages, finally producing merozoites which invade and multiply, via the trophozoite stage, in red blood cells. Eventually, up to 10% of all red cells become infected. (Clinical features of malaria, including fever and chills, anemia and cerebral malaria, are all associated with infected red blood cells, and most current drugs target this stage of the life cycle.) The merozoites in a subset of infected red blood cells then develop into gametocytes. B. When another mosquito bites the infected human, it takes up blood containing gametocytes, which develop into male and female reproductive cells (gametes). These fuse in the insect's gut to form a zygote. The zygote in turn develops into the ookinete, which crosses the wall of the gut and forms a sporozoite-filled oocyst. When the oocyst bursts, the sporozoites move to the mosquito's salivary glands, and the process begins again. Figure and Caption from Wirth, 2002[6]

1.3 The Genome of *P. falciparum*

The *P. falciparum* genome is comprised of approximately 23.3 megabases (Mb), structured into fourteen chromosomes ranging in size from about 0.64 to 3.2 Mb, and contains roughly 5,300 genes. The complete haploid genome sequence of *P. falciparum* was published in 2002, which allowed for new insights into the molecular biology of the parasite, especially genome organization and gene expression.[7] The base composition of (A+T) is atypically high at 80.6% and reaches about 90% in introns and intergenic regions.[7] In comparison, the A+T content of *Saccharomyces cerevisiae* is about 60%.[8] The organization of each individual chromosome is similar, with single copy genes needed for essential cell function and replication in central regions and members of the large, multi-copy, clonally variant gene families found in long tandem arrays within the subtelomeric domains.[9] Adjacent to these multi-copy gene families are conserved telomere associated repeat elements (TAREs), which typically extend about 10-120 kb followed by the actual chromosome ends, which are canonical telomeric heptad repeats maintained by telomerase. (Figure 1.2) The TAREs are separated into five conserved subtelomeric blocks. TARE-1 contains the 7-bp telomeric repeat in a variable number of copies.[10] TARE-2 has sub-blocks of repeats of varying sizes, including 10 copies of a 135 bp repeat, the main element of TARE-2. TARE-3 consists of the Rep20 element, which contains highly variable copies of a 21bp repeat. TARE-4 contains the previously described R-FA3 sequence[9] and short tandem repeats. TARE-5 has no tandem repeats and is found in half of the subtelomeres.[7] The organization and sequence of the TAREs suggest that there is frequent recombination between subtelomeric regions.

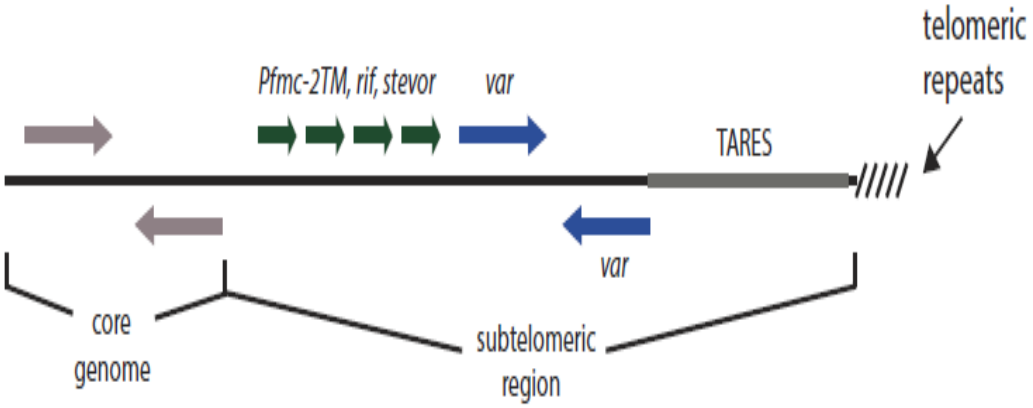


Figure 1.2. The structure of chromosome and subtelomeric domain of *P. falciparum*. The core genome contains primarily single-copy housekeeping genes (gray), while the subtelomeric regions consists of large arrays of variant antigen-encoding genes of the *var* (blue), *rif*, *stevor*, and *Pfmc-2TM* (green) families. TAREs are positioned between the variant gene families and the telomere repeats. Figure and Caption from Calhoun et al., 2017[11]

The subtelomeric domains of *P. falciparum* are unique and integral to understanding parasite biology, as the clonally variant gene families are primary virulence determinants and encode proteins that are exposed on the surface of the infected erythrocyte.[12] This region contains hundreds of genes from several gene families including *var, rifin, stevor, Pfmc-2TM, FIKK* and *ACS*. The most studied family is *var,* the main virulence factor of the disease. It encodes a polymorphic protein whose variant forms are called *Plasmodium falciparum* erythrocyte membrane protein 1 (*Pf*EMP1).[13-15] An adhesion protein, it allows the parasite-infected erythrocyte to adhere to the vasculature of the host, preventing it from being cleared by the spleen. The varied expression of these proteins causes antigenic variation, a mechanism which allows the parasite to avoid antibody mediated clearance and sustain long-term, chronic infections.[5] These multicopy gene families within the subtelomeric domains are found in a specialized chromatin structure that is delineated by the histone modifications H3K9me3 and H3K36me3.[16, 17] These epigenetic marks are known to play a role in mutually exclusive expression (expression of only one gene while remainder of the family is transcriptionally silent) and transcriptional switching (synchronized changes transcriptionally active gene within a multi-copy gene family). H3K9ac is found at the promoter of the single active gene while H3K9me3 is associated with the silent remaining members of the family[18, 19] and H3K36me3 is found at all *var* genes and is important to maintain mutually exclusive expression[20, 21]. In the nucleus, these regions cluster at the nuclear periphery and the homology between regions allows for recombination between the members of the multicopy gene families that are on different chromosomes which facilitates the

generation of diversity.[22] These genes are found to be highly recombinogenic and this creates the hypervariability of the encoded proteins. This hypervariability, along with mutually exclusive expression and transcriptional switching, allows the parasite to evade detection by the host, disrupting the ability of the immune system to produce antibodies to these exposed immunogenic proteins. Given this, these regions are highly important for the study of DNA repair, replication and structural maintenance with an eye towards vaccine development and potential drug targets.

1.4 Understanding DNA Repair of *Plasmodium falciparum*

It is imperative that the genetic information of organisms be functional and available for replication. But the genome is constantly being bombarded by environmental and endogeneous agents that cause a variety of DNA lesions.[23] Endogenous DNA damage can occur spontaneously even under normal physiological conditions because of the inherent instability of chemical bonds in the DNA structure. Some common DNA lesions are pyrimidine dimers, intrastrand crosslinks, deaminated bases, and mispaired bases, but double strand breaks (DSBs) are most deleterious to the cell.[23] To cope with these DNA lesions, eukaryotes have complex networks of DNA repair mechanisms. The DNA repair pathways of *Plasmodium* are Base Excision Repair (BER), Mismatch Repair (MMR), Nucleotide Excision Repair (NER) and Double Strand Break Repair. BER repairs oxidized, deaminated, and alkylated bases and single strand breaks using a Class II AP parasite endonuclease.[24] BER is probably the most used DNA repair pathway in a eukaryotic cell.[25] MMR repairs mismatched bases which primarily result from errors by DNA polymerases during

DNA replication and insertion/deletion loops.[26] NER repairs pyrimidine dimers and intra-strand crosslinks and the proteins involved are highly conserved in the *Plasmodium* genome.[27] Finally, DSBs can be caused by endogenous sources such as reactive oxygen species (ROS) generated by metabolism, transcription or replication fork collapse[28] or experimental sources including ionizing irradiation, chemical mutagens and DNA nucleases. For *Plasmodium* specifically; oxidative damage caused by hemoglobin degradation could be a cause of DSBs. As the parasite grows inside an erythrocyte, it degrades large quantities of hemoglobin in the digestive vacuole, releasing heme. Heme is oxidized causing hydroxyl radicals, which is a strong DNA damaging agent.[29] Antimalarial drugs such as artesunate can cause oxidative DNA damage in mammalian cells[30] and free radicals generated by chloroquine[31] and artemisinin[32] may contribute to DSBs. These assaults on the genome are problematic as a DSB is lethal to the cell unless repaired.

Eukaryotes repair DSBs using two different pathways, either homologous recombination (HR) or non-homologous end joining (NHEJ). While NHEJ can rapidly ligate broken DNA ends, the repair mechanism results in insertions and deletions at the site of ligation and can result in a mutated sequence, which can be deleterious to the organism.[33] HR is a more accurate method of repair since it uses a homologous template from somewhere else in the genome, often the matching allele in a diploid cell, and makes an error-free exact copy. The homologous template must come from a region of the genome with high sequence identity (typically greater than 98%) for the repair to occur.[34] When a DSB is repaired by HR, the DSB is sensed by the Mre11-Rad50-Xrs2 (MRX) complex and the MRX complex binds to the double-

stranded DNA ends and initiates 5' to 3' resection with Sae2. Initial resection is followed by extensive 5' to 3' resection with exonuclease 1 (Exo1) and the Sgs1-Top3-Rmi1-Dna2 (STR-Dna2) complex), this commits the repair process to HR and generates long 3′ ssDNA tails.[28] The amount of resection can vary depending on the species; in *P. falciparum* the A-T rich genome with long regions of repetitive, low complexity sequence may resect extensively to ensure the resected region is unique to undergo HR. The 3' ssDNA ends made by resection are the covered by the ssDNA binding protein Replication Protein-A (RPA). Recombination begins when Rad51 recombinase displaces RPA to form a nucleo-protein filament on the 3' end of ssDNA.[35] The 3' end then invades and anneals to the homologous template, from either a homologous chromosome or sister chromatid. Rad51 promotes strand exchange, which requires ATP hydrolysis activity, forming a displacement loop. Then polymerase δ-dependent DNA synthesis occurs and either synthesis-dependent strand annealing or resolution of the Holliday junctions concludes the repair.[36]

P. falciparum has the homologs of HR proteins such as Rad51[37] and Mre11[38] but bioinformatics analyses have failed to identify any of the key components of the classical NHEJ (C-NHEJ) pathway. The Ku70/80 proteins are thought to be required for C-NHEJ but cannot be found encoded in the genome along with Ligase IV, Artemis and XRCC4, all key components of the C-NHEJ pathway.[7, 39] (Figure 1.3) Interestingly, in a closely related Apicomplexan parasite *Toxoplasma gondii* these key proteins are easily identified and it has been shown that NHEJ is the dominant pathway to repair DNA DSBs in this organism.[40] Despite evolutionary proximity to *T. gondii*, *P. falciparum* has lost its NHEJ machinery suggesting a selective advantage

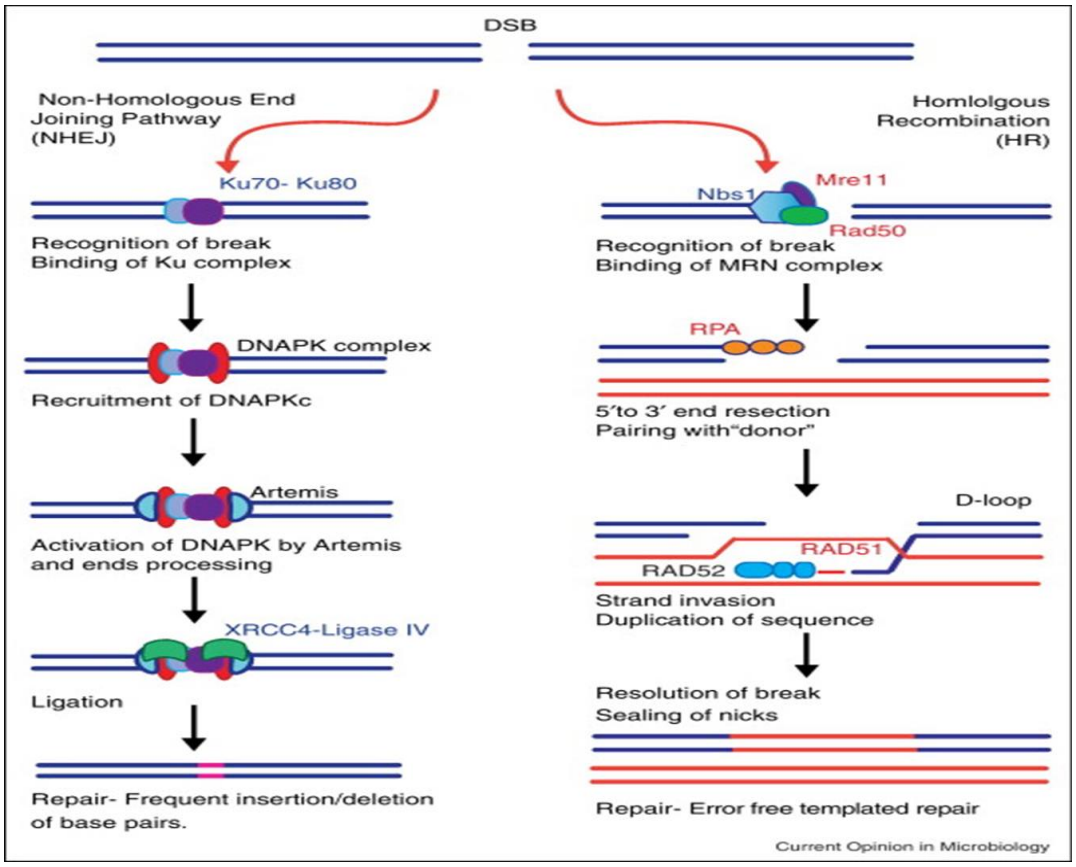


Figure 1.3. Mechanisms of DNA Double Strand Break Repair. Double Strand Breaks (DSBs) are repaired by two major pathways. Homologous Recombination (HR, right) or Nonhomologous End Joining (NHEJ, left). HR can result in different end products, but is defined by the use of a template to guide repair. Pictured is HR leading to a gene conversion. A DNA lesion is recognized by the MRN (MRE11/RAD50/NBS1) complex which generates single stranded DNA by resection. The newly generated single strand is bound by replication protein A (RPA), followed by RAD 52 which acts as a mediator between RPA and RAD51. RAD51 catalyzes the invasion of the single strand to form a displacement loop (D- loop) which is resolved through alternative pathways to result in accurate repair. HR frequently results in a gene conversion event. In contrast, NHEJ does not require a template and instead simply ligates the two free DNA ends together to repair the break. In classical NHEJ, Ku 70/80 sense the DSB and recruit DNA dependent protein kinase complex (DNAPKc) which likely regulates the processing of DNA ends in addition to recruiting the other components of repair including XRCC4, DNA ligase IV, XLF and Artemis. This type of repair frequently results in deletions or insertions at the site of repair (shown as a red insertion in the final product). Components of the repair pathways shown in blue have not been found in the P. falciparum genome; those in red have orthologues identified in the parasite genome. Caption and Figure from Kirkman and Deitsch, 2012[41]

for utilizing one type of repair, HR in this case, and the effect this may have on the genome. Examining a phylogenetic tree of the Alveolata superphylum, the ability to perform C-NHEJ, which is defined by identifiable Ku proteins, is present for some organisms but not others. Using NCBI blast with the yeast Ku70 sequence as template, the Ku proteins and Ligase IV can be identified for the Ciliates *Perkinsus marinus* and *Paramecium tetrauralia*. But when searching within the Apicomplexan parasites the orthologs of the Ku proteins can only be identified in the Coccidian branch which includes *Toxoplasma, Eimeria* and *Neospora*. In contrast, a search for the key proteins of HR (Rad51 and Mre11) can easily be identified for all the Alveolates.[12] *(*Figure 1.4) Though *Plasmodium* may lack C-NHEJ, evidence suggests that *Plasmodium* is able to utilize an alternative non-homologous end joining pathway (A-NHEJ) similar to microhomology-mediated end joining (MMEJ).[39, 42] This pathway uses a non-processive DNA polymerase to create microhomology at the site of the DSB which then allows for ligation and repair; this is specifically described as synthesis-dependent microhomology mediated end joining (SD-MMEJ).[43] The characteristics of this pathway are single base pair insertions or deletions at or near the site of the break and short segments of sequence that appear to be template from regions very close to the break site.[43] A possible polymerase involved was identified encoded in the *P. falciparum* genome that contains the Pol A superfamily domain (PF3D7_0625300) found in pol theta, which is the polymerase used in *Drosophila,* but the full function of this polymerase has yet to be explored. Ligase III is also required for MMEJ but is not found in *Plasmodium*[42], though Ligase III may be replaceable by Ligase I[44], which *Plasmodium* does encode. Regardless, the rate of repair of this

pathway was shown to be inefficient in *Plasmodium* suggesting that this pathway is not an often utilized method of repair.[39]

Because the *P. falciparum* genome is haploid, the question remains, how does the parasite repair DNA double strand breaks if C-NHEJ is not present and A-NHEJ is inefficient? Evidence suggests HR is the repair pathway that is used exclusively but how can this occur when there is no homologous template for repair available? The parasite may be able to utilize HR during the trophozoite and schizont stage (before segmentation) when active DNA replication allows access to sister chromatids, thus a homologous sequence to the site of the DSB could be utilized. Examination of one of the multicopy gene families, *var,* indicates that these genes undergo gene conversion events to promote antigenic variation though these genes are more than 2% divergent from each other (the usual threshold of templates for HR repair).[12] It has been hypothesized that HR can occur at these genes located in the subtelomeric region and internal clusters because of the unique histone modifications and chromatin structure at these sites.[12] In contrast, inducing DSBs at non-*var* regions of the genome where the sequence surrounding the site of the break diverged more than 2% from any possible template for repair resulted in exclusively A-NHEJ being used for repair and no evidence of gene conversion events were found.[39] In the subtelomeric regions of the genome, DSBs can also be repaired by a pathway called de novo telomere addition or "telomere healing" which creates a functional telomere but with a loss of the region of the chromosome between the original telomere and the site of the DSB.[45] This process has been described for *P. falciparum*[46].

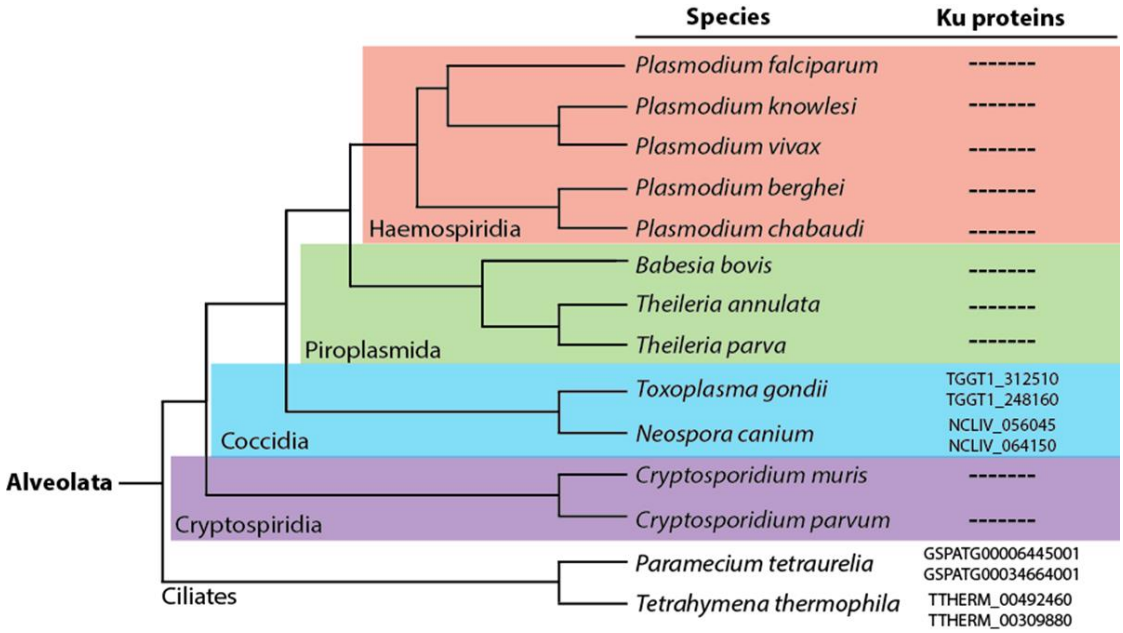


Figure 1.4. Phylogenetic tree of the Apicomplexan lineage. The different colors highlight several groups of obligate, Apicomplexan parasites. The Haemosporidia (orange) include parasites that cause malaria in humans (*P. falciparum* and *P. vivax*), nonhuman primates (*P. knowlesi*) and rodents (*P. berghei* and *P. chabaudi*). The Piroplasmida (green) are parasites transmitted by ticks that infect either red blood cells (*Babesia bovis*) or white blood cells (*T. annulata* and *T. parva*). The Coccidia (blue) are cyst forming parasites that do not undergo antigenic variation. The Cryptospiridia (purple) also form cysts and have a highly reduced genome. The ciliates, as exemplified by *P. tetraurelia* and *T. thermophila*, are nonparasitic, free living organisms. The ability of the organisms to utilize the C-NHEJ pathway for DSB repair is inferred by the presence or absence of proteins of the Ku family. Note that, of the Apicomplexans, only the Coccidia have retained these genes within their genomes. The annotations numbers for Ku70/80 are given for both the Coccidia and the Ciliates. Caption and Figure from Kirkman et al. 2014[12].

We also have identified "telomere healing" as a pathway for DSB repair and its competition with HR for DSB repair within subtelomeric regions. This appears to depend solely on the DNA sequences surrounding the site where the DSB occurs, which will be discussed in more detail in Chapter 3.

Beyond this method of repair, it is not fully understood how the parasite repairs regions of its genome that do not have a unique chromatin structure when the parasite is haploid and not undergoing DNA replication. The malaria parasite is able to withstand DNA damage, assumedly repairing DSBs and maintaining infections within human hosts, but how exactly this is managed is unclear. A possibility is that RNA-templated repair is used when a DSB occurs and the parasite has no DNA homologous template for repair. A study using synthetic RNA oligonucleotides done in *Saccharomyces cerevisiae* showed that RNA could provide a template for DNA synthesis during repair of a DSB.[47] This finding was corroborated in a human kidney cell model using synthetic oligonucleotides.[48] However, in 2014 Keskin et al. found that yeast could use transcript-RNA for direct DSB repair, indicating that this could happen naturally *in vivo* using endogenous RNA transcripts. The RNA transcript templated DSB repair occurred at higher frequency when the RNA was derived from the same locus as the break but also occurred when the homologous RNA was from distant parts of the genome.[49] Experiments suggest that the Rad52 protein can mediate binding of complementary RNA and DNA strands[50] but *P. falciparum* lacks this protein[28]. However, in Rad52 null yeast mutants DSB repair with transcript RNA was not abolished, indicating the existence of a Rad52 independent pathway.[50] In addition the recombinase Rad51 can promote RNA-DNA hybridization, strand exchange *in*

vitro, and R-loop formation *in vivo*.[51] Rad51 is also preferentially recruited to sites of transcriptionally active chromatin that have DSBs.[52] Reviewing these findings, it is possible that transcript RNA plays a role in DSB repair in *P. falciparum* mediated by the Rad51 protein, in which RNA templated repair occurs preferentially at actively transcribed genes but can also occur in *trans* using a transcript RNA from elsewhere in the genome. The ability of the *Plasmodium genome* to repair itself is essential as the immune system of the host generates a response to clear infected erythrocytes that can causes DNA damage to the parasite.[28, 29]

1.5 The Immune Response to *Plasmodium*

A quick and robust innate immune response is essential to limiting the initial phase of malaria infection, controlling parasitemia and increasing odds of host survival so that an adaptive immune response to clear the infection can be mounted.[3, 5] A pro-inflammatory response is triggered by malaria pathogen-associated molecular patterns (PAMPs) such as plasmodial DNA and RNA, haemozoin, and GPI anchors, that are detected by macrophages and dendritic cells which produce IFN-γ, a cytokine crucial for parasite control and clearance.[4] Macrophage and dendritic cells are phagocytic, antigen presenting cells and are often the first to encounter a pathogen and activate other immune cells through antigen presentation or cytokine production.[3, 5] Natural killer cells, after activation by cytokines released from dendritic cells, are potent producers of IFN-γ, which activates macrophages and continues the positive feedback loop of cytokine production and parasiticidal molecules that propels a pro-inflammatory immune response.[3, 53] Activated macrophages produce reactive oxygen

species (ROS) such as nitric oxide and superoxide which can cause oxidative DNA damage or kill infected erythrocytes.[54] There is growing evidence for cytotoxic killing of parasitized erythrocytes by natural killer cells, by release of their granzymes, which contain serine proteases.[53] However these cells kill parasites via protein cleavage, therefore, the main source of DNA damage from immune cells comes from macrophage produced ROS. The human immune response creates an inflammatory environment designed to isolate, damage and kill malaria parasites.

1.6 Genetic Mutations and Drug Resistance

Repair of DNA damage caused by the immune response, anti-malarials, or external sources allows the parasite to introduce mutations into the genome which can be beneficial for genetic variation and developing resistance to drugs.[55, 56] The resistance of *P. falciparum* to several classes of antimalarial drugs, such as chloroquine, pyrimethamine, sulfadoxine and artemisinin, often first emerges in Southeast Asia and South America and then expands globally.[55, 57] These drug resistance phenotypes are correlated with DNA polymorphisms, usually non-synonymous mutations in key drug resistance genes.[57] The rapid emergence of these anti-malarial drug resistance phenotypes is thought to be linked to the increased mutation ability of some parasite strains (mutator phenotype), called accelerated resistance to multiple drugs (ARMD).[58] Defective DNA excision repair pathways such as NER and MMR have been linked to drug resistant phenotypes.[27, 59, 60] A study looking at DNA repair in ARMD and sensitive parasites found that two strains of ARMD parasites had little MMR activity as compared to a sensitive strain, indicating

that the ARMD parasites had a defective MMR pathway.[59] This suggests that parasites with defective MMR are selected for under drug pressure and MMR is an underlying mechanism for the generation of antimalarial drug resistance.[59] After experimental exposure to MMS, non-ARMD parasites upregulated MMR and NER repair machinery genes while ARMD parasites showed no induction. Polymorphisms were found in 18 DNA repair genes including MRE11, which could lead to the impaired DNA repair function observed.[55] Defects in DNA repair machinery were found to mediate mutator phenotypes in two other parasites; *T. gondii*[61] and *T. brucei*[62]. Therefore, DNA repair pathways play a crucial role in the mutator phenotype indicating a direct link between a parasite's ability to repair DNA and drug resistance.

1.7 Thesis Objective

The field of DNA repair in *Plasmodium falciparum* still presents many unanswered questions about the mechanisms of DNA repair pathways and the importance of the proteins involved in these pathways. We attempt to answer some of those questions by examining DSB repair in the subtelomeric domains of the genome and evaluating the essentiality of a major HR repair protein. In Chapter 2 we investigate the difference in sensitivity to DNA damage at each parasite life cycle stage and how the ability to perform HR can influence damage tolerance variably at different parasite life cycle stages. This work adds to previous research examining Rad51 and the parasite's ability to compensate for loss of HR activity with other repair mechanisms. In Chapter 3 we analyze DSB repair at the subtelomeric domains of the

genome using irradiation and discover a competitive balance between "telomere healing" and HR to repair chromosome ends. This works indicates that both pathways are used to repair breaks in this genomic region and the choice depends on the sequence surrounding the break site. In Chapter 4, these findings are summarized and the broader implications of this research within the field are discussed. Overall, we aim to add to the knowledge base of DSB repair in *P. falciparum*, particularly how HR, the major repair pathway, interchanges with alternative repair mechanisms to stabilize the genome of a haploid parasite.

References

1. WHO. 186 (World Health Organization, Geneva, 2016).
2. Haque, U. et al. Malaria burden and control in Bangladesh and prospects for elimination: an epidemiological and economic assessment. *Lancet Glob Health* **2**, e98-105 (2014).
3. Stevenson, M.M. & Riley, E.M. Innate immunity to malaria. *Nat Rev Immunol* **4**, 169-80 (2004).
4. Gazzinelli, R.T., Kalantari, P., Fitzgerald, K.A. & Golenbock, D.T. Innate sensing of malaria parasites. *Nat Rev Immunol* **14**, 744-57 (2014).
5. Miller, L.H., Baruch, D.I., Marsh, K. & Doumbo, O.K. The pathogenic basis of malaria. *Nature* **415**, 673-9 (2002).
6. Wirth, D.F. Biological revelations. *Nature* **419**, 495-6 (2002).
7. Gardner, M.J. et al. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419**, 498-511 (2002).
8. Basile, W., Sachenkova, O., Light, S. & Elofsson, A. High GC content causes orphan proteins to be intrinsically disordered. *PLoS Comput Biol* **13**, e1005375 (2017).
9. de Bruin, D., Lanzer, M. & Ravetch, J.V. The polymorphic subtelomeric regions of Plasmodium falciparum chromosomes contain arrays of repetitive sequence elements. *Proc Natl Acad Sci U S A* **91**, 619-23 (1994).
10. Vernick, K.D. & McCutchan, T.F. Sequence and structure of a Plasmodium falciparum telomere. *Mol Biochem Parasitol* **28**, 85-94 (1988).
11. Calhoun, S.F. et al. Chromosome End Repair and Genome Stability in Plasmodium falciparum. *MBio* **8** (2017).
12. Kirkman, L.A. & Deitsch, K.W. Recombination and Diversification of the Variant Antigen Encoding Genes in the Malaria Parasite Plasmodium falciparum. *Microbiol Spectr* **2** (2014).
13. Baruch, D.I. et al. Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. *Cell* **82**, 77-87 (1995).
14. Su, X.Z. et al. The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes. *Cell* **82**, 89-100 (1995).
15. Smith, J.D. et al. Switches in expression of Plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* **82**, 101-10 (1995).
16. Lopez-Rubio, J.J., Mancio-Silva, L. & Scherf, A. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe* **5**, 179-90 (2009).
17. Flueck, C. et al. Plasmodium falciparum heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. *PLoS Pathog* **5**, e1000569 (2009).
18. Chookajorn, T. et al. Epigenetic memory at malaria virulence genes. *Proc Natl Acad Sci U S A* **104**, 899-902 (2007).
19. Lopez-Rubio, J.J. et al. 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Mol Microbiol* **66**, 1296-305 (2007).
20. Jiang, L. et al. PfSETvs methylation of histone H3K36 represses virulence genes in Plasmodium falciparum. *Nature* **499**, 223-7 (2013).

21. Ukaegbu, U.E. et al. Recruitment of PfSET2 by RNA polymerase II to variant antigen encoding loci contributes to antigenic variation in P. falciparum. *PLoS Pathog* **10**, e1003854 (2014).
22. Figueiredo, L.M., Freitas-Junior, L.H., Bottius, E., Olivo-Marin, J.C. & Scherf, A. A central role for Plasmodium falciparum subtelomeric regions in spatial positioning and telomere length regulation. *EMBO J* **21**, 815-24 (2002).
23. Altieri, F., Grillo, C., Maceroni, M. & Chichiarelli, S. DNA damage and repair: from molecular mechanisms to health implications. *Antioxid Redox Signal* **10**, 891-937 (2008).
24. Haltiwanger, B.M. et al. DNA base excision repair in human malaria parasites is predominantly by a long-patch pathway. *Biochemistry* **39**, 763-72 (2000).
25. Morita, R. et al. Molecular mechanisms of the whole DNA repair system: a comparison of bacterial and eukaryotic systems. *J Nucleic Acids* **2010**, 179594 (2010).
26. Tarique, M., Ahmad, M., Chauhan, M. & Tuteja, R. Genome Wide In silico Analysis of the Mismatch Repair Components of Plasmodium falciparum and Their Comparison with Human Host. *Front Microbiol* **8**, 130 (2017).
27. Tajedin, L., Anwar, M., Gupta, D. & Tuteja, R. Comparative insight into nucleotide excision repair components of Plasmodium falciparum. *DNA Repair (Amst)* **28**, 60-72 (2015).
28. Lee, A.H., Symington, L.S. & Fidock, D.A. DNA repair mechanisms and their biological roles in the malaria parasite Plasmodium falciparum. *Microbiol Mol Biol Rev* **78**, 469-86 (2014).
29. Atamna, H. & Ginsburg, H. Origin of reactive oxygen species in erythrocytes infected with Plasmodium falciparum. *Mol Biochem Parasitol* **61**, 231-41 (1993).
30. Berdelle, N., Nikolova, T., Quiros, S., Efferth, T. & Kaina, B. Artesunate induces oxidative DNA damage, sustained DNA double-strand breaks, and the ATM/ATR damage response in cancer cells. *Mol Cancer Ther* **10**, 2224-33 (2011).
31. Radfar, A., Diez, A. & Bautista, J.M. Chloroquine mediates specific proteome oxidative damage across the erythrocytic cycle of resistant Plasmodium falciparum. *Free Radic Biol Med* **44**, 2034-42 (2008).
32. O'Neill, P.M., Barton, V.E. & Ward, S.A. The molecular mechanism of action of artemisinin--the debate continues. *Molecules* **15**, 1705-21 (2010).
33. Le Guen, T., Ragu, S., Guirouilh-Barbat, J. & Lopez, B.S. Role of the double-strand break repair pathway in the maintenance of genomic stability. *Mol Cell Oncol* **2**, e968020 (2015).
34. Larocque, J.R. & Jasin, M. Mechanisms of recombination between diverged sequences in wild-type and BLM-deficient mouse and human cells. *Mol Cell Biol* **30**, 1887-97 (2010).
35. Sugiyama, T., Zaitseva, E.M. & Kowalczykowski, S.C. A single-stranded DNA-binding protein is needed for efficient presynaptic complex formation by the Saccharomyces cerevisiae Rad51 protein. *J Biol Chem* **272**, 7940-5 (1997).
36. Symington, L.S. Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol Mol Biol Rev* **66**, 630-70, table of contents (2002).
37. Bhattacharyya, M.K. & Kumar, N. Identification and molecular characterisation of DNA damaging agent induced expression of Plasmodium falciparum recombination protein PfRad51. *Int J Parasitol* **33**, 1385-92 (2003).
38. Badugu, S.B. et al. Identification of Plasmodium falciparum DNA Repair Protein Mre11 with an Evolutionarily Conserved Nuclease Function. *PLoS One* **10**, e0125358 (2015).

39. Kirkman, L.A., Lawrence, E.A. & Deitsch, K.W. Malaria parasites utilize both homologous recombination and alternative end joining pathways to maintain genome integrity. *Nucleic Acids Res* **42**, 370-9 (2014).
40. Smolarz, B., Wilczynski, J. & Nowakowska, D. DNA repair mechanisms and Toxoplasma gondii infection. *Arch Microbiol* **196**, 1-8 (2014).
41. Kirkman, L.A. & Deitsch, K.W. Antigenic variation and the generation of diversity in malaria parasites. *Curr.Opin.Microbiol.* **15**, 456-462 (2012).
42. Singer, M. et al. Zinc finger nuclease-based double-strand breaks attenuate malaria parasites and reveal rare microhomology-mediated end joining. *Genome Biol* **16**, 249 (2015).
43. Yu, A.M. & McVey, M. Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res* **38**, 5706-17 (2010).
44. Paul, K. et al. DNA ligases I and III cooperate in alternative non-homologous end-joining in vertebrates. *PLoS One* **8**, e59505 (2013).
45. Pennaneach, V., Putnam, C.D. & Kolodner, R.D. Chromosome healing by de novo telomere addition in Saccharomyces cerevisiae. *Mol Microbiol* **59**, 1357-68 (2006).
46. Scherf, A. & Mattei, D. Cloning and characterization of chromosome breakpoints of Plasmodium falciparum: breakage and new telomere formation occurs frequently and randomly in subtelomeric genes. *Nucleic Acids Res* **20**, 1491-6 (1992).
47. Storici, F., Bebenek, K., Kunkel, T.A., Gordenin, D.A. & Resnick, M.A. RNA-templated DNA repair. *Nature* **447**, 338-41 (2007).
48. Shen, Y. et al. RNA-driven genetic changes in bacteria and in human cells. *Mutat Res* **717**, 91-8 (2011).
49. Keskin, H. et al. Transcript-RNA-templated DNA recombination and repair. *Nature* **515**, 436-9 (2014).
50. Meers, C., Keskin, H. & Storici, F. DNA repair by RNA: Templated, or not templated, that is the question. *DNA Repair (Amst)* **44**, 17-21 (2016).
51. Wahba, L., Gore, S.K. & Koshland, D. The homologous recombination machinery modulates the formation of RNA-DNA hybrids and associated chromosome instability. *Elife* **2**, e00505 (2013).
52. Aymard, F. et al. Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat Struct Mol Biol* **21**, 366-74 (2014).
53. Wolf, A.S., Sherratt, S. & Riley, E.M. NK Cells: Uncertain Allies against Malaria. *Front Immunol* **8**, 212 (2017).
54. Fritsche, G., Larcher, C., Schennach, H. & Weiss, G. Regulatory interactions between iron and nitric oxide metabolism for immune defense against Plasmodium falciparum infection. *J Infect Dis* **183**, 1388-94 (2001).
55. Gupta, D.K., Patra, A.T., Zhu, L., Gupta, A.P. & Bozdech, Z. DNA damage regulation and its role in drug-related phenotypes in the malaria parasites. *Sci Rep* **6**, 23603 (2016).
56. Aravind, L., Iyer, L.M., Wellems, T.E. & Miller, L.H. Plasmodium biology: genomic gleanings. *Cell* **115**, 771-85 (2003).
57. Ashley, E.A. et al. Spread of artemisinin resistance in Plasmodium falciparum malaria. *N Engl J Med* **371**, 411-23 (2014).
58. Rathod, P.K., McErlean, T. & Lee, P.C. Variations in frequencies of drug resistance in Plasmodium falciparum. *Proc Natl Acad Sci U S A* **94**, 9389-93 (1997).
59. Castellini, M.A. et al. Malaria drug resistance is associated with defective DNA mismatch repair. *Mol Biochem Parasitol* **177**, 143-7 (2011).

60. Trotta, R.F., Brown, M.L., Terrell, J.C. & Geyer, J.A. Defective DNA repair as a potential mechanism for the rapid development of drug resistance in Plasmodium falciparum. *Biochemistry* **43**, 4885-91 (2004).
61. Garrison, E.M. & Arrizabalaga, G. Disruption of a mitochondrial MutS DNA repair enzyme homologue confers drug resistance in the parasite Toxoplasma gondii. *Mol Microbiol* **72**, 425-41 (2009).
62. Augusto-Pinto, L., Teixeira, S.M., Pena, S.D. & Machado, C.R. Single-nucleotide polymorphisms of the Trypanosoma cruzi MSH2 gene support the existence of three phylogenetic lineages presenting differences in mismatch-repair efficiency. *Genetics* **164**, 117-26 (2003).

Chapter 2: The Essentiality of Rad51-Mediated Repair

2.1 Introduction

During its lifecycle, the malaria parasite is faced with many factors that can cause DNA damage, including products of hemoglobin digestion, the immune response of the host, errors during DNA replication and oxidative damage from antimalarial drugs.[1-3] In order to survive, these parasites must be able to efficiently repair DNA damage as one unrepaired double strand break (DSB) can cause the death of a unicellular organism.[4] The erythrocytic stage of *P. falciparum* has a unique DNA repair dilemma, since the organism is haploid; they lack homologous sequences as templates for HR repair but also lack the ability to perform C-NHEJ. An alternative NHEJ pathway has been characterized but this method appears inefficient.[5, 6] Therefore, how the parasite repairs it's DSBs and maintains the structural integrity of its genome is enigmatic. As such targeting the HR pathway, the main DSB repair method for *Plasmodium,* may provide a key to controlling the disease.

Rad51 is the RecA recombinase in eukaryotes and plays a major role in HR-mediated DSB repair after a form of DNA damage has occurred.[7] When a DSB occurs the broken DNA ends are nucleolytically processed by several players, including the MRE11-RAD50-XRS2/NBS1 complex, DNA2, EXO1, and CTIP [8], which exposes single-stranded DNA (ssDNA) which is rapidly covered by the ssDNA binding protein Replication Protein-A (RPA). Recombination begins when Rad51 recombinase displaces RPA to form a nucleo-protein filament on the 3' end of ssDNA.[9] The 3' end then invades and anneals to the homologous template, from

either a homologous chromosome or sister chromatid. Rad51 promotes strand exchange, which requires ATP hydrolysis activity, then DNA synthesis occurs and resolution of the Holliday junctions concludes the repair.[7] The Rad51 protein has been identified and kinetically characterized in *P. falciparum* by Bhattacharyya et al.[10, 11] The PfRad51 protein has ssDNA-dependent ATPase activity and exhibits strand exchange activity dependent on ATP *in vitro.*[11] Replication protein-A (RPA) has also been identified in *P. falciparum*.[12] Though HR protein orthologs have been found in *Plasmodium* the exact mechanisms and limitations of HR in the parasite are not well understood. A knockout of *rad51* in *Saccharomyces cerevisiae* renders cells hypersensitive to DNA damaging agents[13-15] while in mice Rad51 knockout leads to embryonic lethality[16, 17]. Mutations of Rad51 have been characterized in several organisms. The conserved Walker A and Walker B motifs bind ATP.[18] A mutation of a highly conserved lysine (K133 in human Rad51) in the Walker A motif is defective for ATP hydrolysis.[19] The HsRad51^{K133R} mutant is deficient for ATPase function and demonstrates a dominant negative phenotype that inhibits WT hRad51 ATPase activity[20 19, 21] Cells expressing HsRad51^{K133R} and WT hRad51 are more sensitive to DNA damage than cells expressing only WT HsRad51 [19, 20] and expression of the mutant inhibited homologous recombination and intrachromosomal recombination.[20] A dominant negative mutant (PfRad51^{K143R}) was first described in a mouse malaria model where parasites harboring this mutation were found to have reduced survival after exposure to mutagenesis.[18] We expressed the PfRad51 ^{K143R} in *P falciparum in vitro* to examine the efficiency of DSB repair in *P. falciparum in vitro.*

The malaria parasite has three developmental stages in the asexual erythrocytic life cycle stage: ring, trophozoite and schizont. A haploid (1n) merozoite invades an erythrocyte and develops as a ring form from 0 to about 24 hours post-invasion. The ring progresses to a trophozoite (1n to 2n), and at mid trophozoite stage the parasite undergoes DNA synthesis. The trophozoite then becomes a schizont, undergoing repeated mitosis and nuclear division and producing up to 24 nuclei.[1, 22] The DNA replication occurring during the trophozoite and schizont stages create access to sister chromatids which allows for efficient homologous recombination. This implies that parasites can repair DSBs with varying efficiency depending on which stage of the cycle the DSB is created. If DNA repair is primarily dependent on HR as proposed, then trophozoite parasites should be significantly less sensitive to DSBs since they have access to a homologous template for repair. Furthermore, because of this, repair efficiency in trophozoites should be more sensitive to Rad51DN then repair in rings and segmented schizonts.

DSB repair mediated by Rad51 in *P.falciparum* has been studied using Methyl Methanesulfonate (MMS) [18] but in this work we expose erythrocytic stage parasites to ionizing radiation to create broad spectrum DNA damage. Ionizing radiation such as X-rays create significant amounts of clustered DNA damage, including DSBs.[23] DNA lesions caused by ionizing radiation are chemically identical to those created by ROS[24] (an *in vivo* cause of DSBs for parasites) while MMS alkylates DNA bases and induces DSBs. [25]

Here we examine how *P. falciparum* parasites repair DSBs at different stages of their replicative cycle and how the availability of homologous templates affects the efficiency of irradiation mediated DNA damage repair. Further examining the components of the HR pathway, we created a PfRad51^{K143R} parasite line and analyzed the irradiation resistance as compared to WT. These experiments help elucidate the role of Rad51, a key protein in HR, and the functionality of HR at different stages of the parasite asexual replication cycle.

2.2 Results

2.2.1 The irradiation tolerance of the malaria parasite.

In determining irradiation tolerance of the malaria parasite in standard in vitro conditions, initial dosage levels were based on studies done by Waki et al. where parasites were exposed to gamma and X-ray irradiation.[26, 27] *P. falciparum* dose response curves using gamma irradiation showed that parasites could survive 200 Gray (Gy) of irradiation, though with less than .01% of the original culture remaining.[26] Another study also confirmed that parasite death and recovery was dependent on the dosage of gamma irradiation.[28] The Waki et al. study using X-ray irradiation exposed parasites up to 100 Gy and found this dosage caused a 1% survival of the culture.[27] After recovery from either X-ray or gamma irradiation, the parasites displayed similar growth curves as the non-irradiated parasites. Using this data we exposed our 3D7 parasites to up to 125 Gray X-ray irradiation and found our results corroborated the irradiation tolerance seen by Waki et al., using days to 2% parasitemia as our indication of dosage response. (Table 2.1) 100Gy was the chosen

dose for our subsequent assays to produce a significant but survivable knockdown of parasite growth. Interestingly, neither we nor the previous work have identified a lethal dose of irradiation for *P. falciparum.*[26, 27] But *P. falciparum* was found to recover after 600 Gy (60 kilorads) of gamma irradiation; a lethal dose has not yet been determined.[28]

Table 2.1. Irradiation Dose Response of 3D7 parasites. Mixed stage parasites were irradiated at 0.5% and cultured until 2% parasitemia was reached. Data from one representative experiment.

Days from Irradiation to 2% Parasitemia	
Irradiation Dose	**Days**
Control (none)	1
3Gy	1
5Gy	1
10Gy	2
50Gy	2
80Gy	7
100Gy	11
125Gy	16

2.2.2 Differences in Irradiation sensitivity of parasites at each life cycle stage. Sensitivity to irradiation induced damage is partially dependent on the ability of the cell to repair damage to its DNA. This dependence is exemplified by observation that genes involved in DNA repair are required for toleration of irradiation damage in yeast[29] and in mammalian cells[30]. This correlation was also indicated in the previous irradiation sensitivity work done with *P. falciparum* in which lethal irradiation damage was found mainly in DNA.[27] These studies suggested that irradiation sensitivity was correlated with the cycle of DNA synthesis in the parasites.[26, 27]

Given the haploid state of the parasite and its dependence on HR for the repair of DSBs, we hypothesized that the parasite's ability to repair its DNA changes as it progresses through each stage of asexual replication depending on whether or not DNA synthesis is occurring. Thus we expected that ring and segmented schizont stage parasites would have similar X-ray irradiation sensitivity, while trophozoites undergoing active DNA replication, would be less sensitive.

To obtain schizonts, parasites were synchronized using 5% (w/v) sorbitol, as described in methods, 3 days and 1 day prior to initiating the experiment. The day of irradiation, 40/70% Percoll-Sorbitol gradient centrifugation was used to obtain only late trophozoites and schizonts. These were then treated with E64 to prevent egress from the erythrocyte[31, 32] and parasites were irradiated when fully segmented. Synchronized WT/DHOD or 3D7 parasites were seeded at 0.5% before irradiation and followed post irradiation until the culture reached 2% parasitemia, a marker of return

to normal growth. (Fig. 2.1) In comparison, Waki et al. synchronized cultures with sorbitol three times at ring stage and then parasites were irradiated at 42 and 48 h.p.i.[27]

As shown in Figure 2.1A, WT/DHOD rings and WT/DHOD schizonts had similar recovery rates at 80 Gy exposure with 9 days (+/- 2.39) for rings and 8 days (+/- 1.64) for schizonts. Similarly, at 100 Gy exposure WT/DHOD rings recovered to normal growth within 16 days (+/- 2.44) and 3D7 schizonts at 15 days (+/- 3.71). Both WT/DHOD and 3D7 parasites were used for the schizont irradiation sensitivity assays as both strains have similar post irradiation growth. (Fig 2.2) In comparison, trophozoite stage parasites with active DNA replication recovered significantly faster, in 4 days (+/- 2.34) at 80 Gy exposure and 9 days (+/- 0.88) at 100 Gy exposure. Tolerance for irradiation peaks during DNA synthesis and decreases as the parasites complete schizogony and reinvade becoming ring stage parasites.[26, 27]In the previous study, Waki et al. irradiated schizonts at 42 h.p.i and next at 48 hours when those schizonts had ruptured, missing the window when all schizonts are fully segmented (between 44-48 h.p.i) [33] which is the time point at which we irradiated the parasites. This difference in irradiation time point for schizonts would explain why previously Waki et al. reported schizont irradiation tolerance similar to trophozoites while we find tolerance to be closer to that of rings.

Parasites exposed to 100Gy of irradiation recovered more slowly than at 80Gy exposure for rings, trophozoites and schizonts. Previous work has also shown that parasites tolerate irradiation in a dose dependent manner.[28] At 80Gy the difference in irradiation sensitivity between rings/schizonts and trophozoites is significant and the

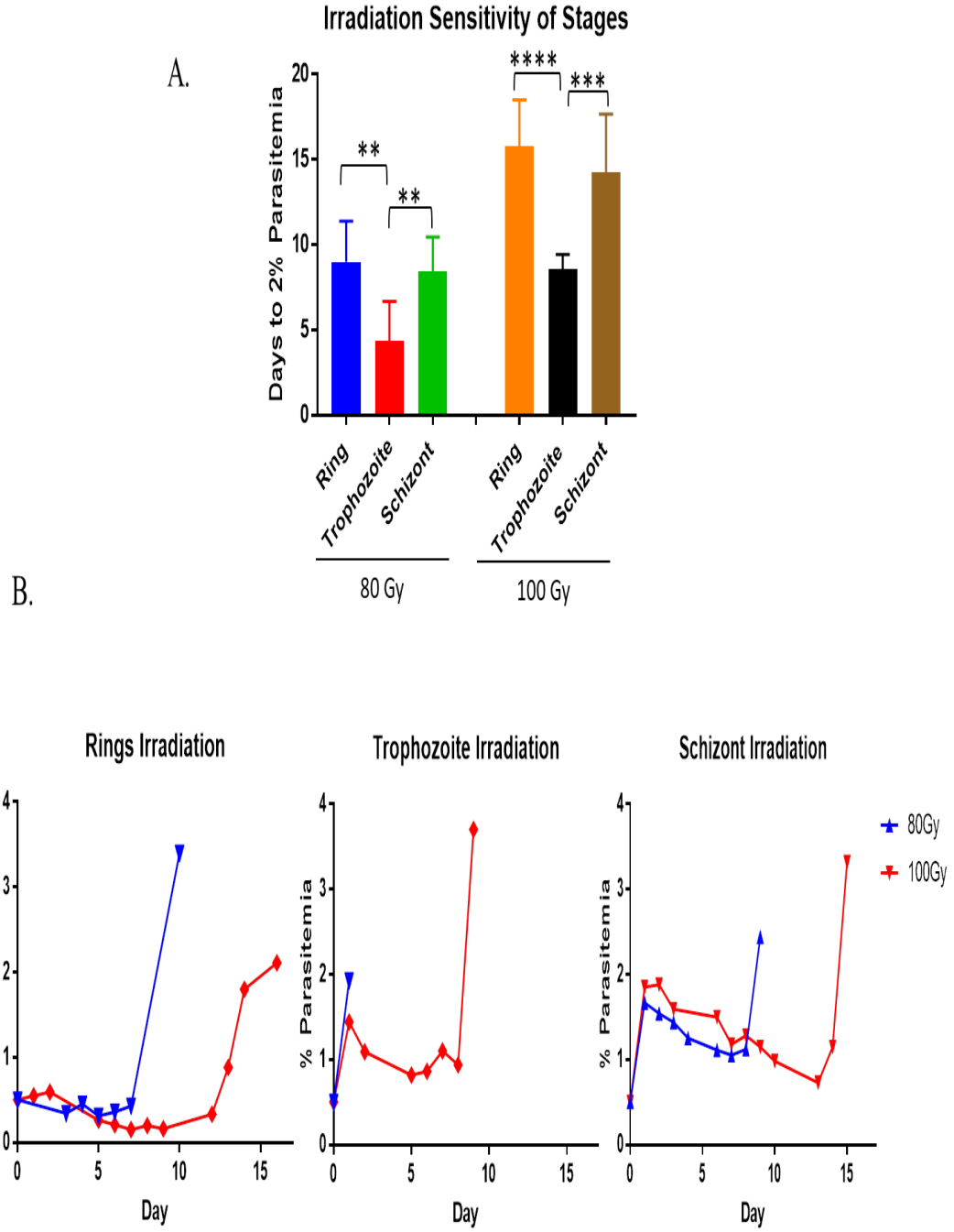


Figure 2.1. Differences in Irradiation sensitivity of parasites at each life stage. A. The differences in days to 2% parasitemia between ring, trophozoites and schizonts at 80 and 100 Gy. Data is from three independent experiments with three replicates in each. **- P<0.01, ***- P<0.001, ****- P<0.0001 by Unpaired t-test with Welch's correction. B. Growth curves at 80 and 100 Gy for each life stage, data plotted is from one representative experiment. Parasites were synchronized and irradiated as described above.

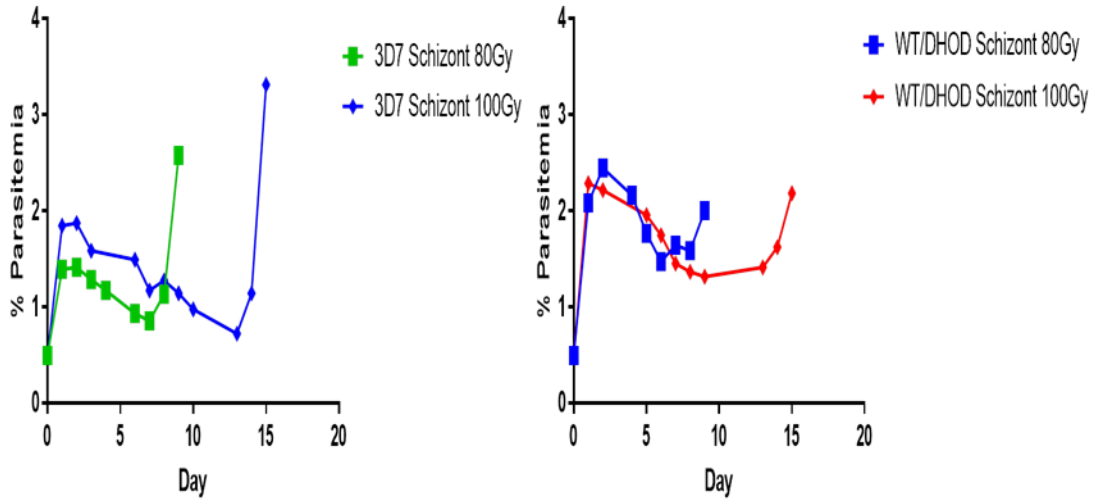


Figure 2.2. Comparison of Irradiation Sensitivity between 3D7 and WT/DHOD schizonts. Growth curves after irradiation for 3D7 and WT/DHOD at 80 and 100 Gy, as both strains are used for schizont irradiation sensitivity data.

difference at 100Gy between ring/schizonts and trophozoites is highly significant. The difference in irradiation sensitivity between rings/schizonts and trophozoites is likely due to the fact that during DNA replication access to sister chromatids are possible before mitosis is complete.[1] This allows the trophozoite parasite to repair DSBs more readily by HR because of homologous template access. Looking at the growth curves (Figure 2.1B), the different growth patterns between each of the three life stages can be visualized. Rings do not increase in parasitemia after irradiation and parasitemia is fairly consistent until recovery of normal growth, this is most apparent at 100 Gy exposure. Post irradiation, trophozoite cultures demonstrate an increase in parasitemia of about 4-fold, which decreases as parasites die, then rebounds after normal growth is recovered. Schizonts have a similar growth pattern to trophozoites but take longer to return to normal growth. In contrast, unirradiated parasites multiplied by about 8 fold per 48 hour life cycle after schizont rupture and reinvasion with no decrease in parasitemia, which is corroborated by previous work[34]. (data not shown) (Table 2.1) Irradiation sensitivity assays at 50 Gy (data not shown) indicate only a minor difference in growth for rings and trophozoites.

2.2.3 Creation of the R51 DN parasite line.

Rad51 is the highly conserved RecA recombinase in eukaryotes and it plays a central role during HR mediated DSB repair, especially in homology search and strand pairing.[9, 35] The conserved Walker A and Walker B motifs bind ATP[18] and the ATP hydrolysis activity of these motifs is necessary for its strand exchange capability.[7] A mutation in the Walker A motif at a highly conserved lysine (K133R) in the human enzyme was found to be deficient in ATPase function, disrupting the function of the

protein.[19, 21] Pf Rad 51 has been initially characterized.[10, 11] and the analogous lysine residue (K143) within the Walker A motif (Figure 2.3A) was identified in PfRad51 by Roy et al.[18] A dominant negative PfRad51^{K143R} was created and the recombinant protein displayed miniscule ssDNA-dependent ATP hydrolysis.[18] Figure 2.3A shows a schematic of the PfRad51^{K143R} with the red line indicating the single nucleotide polymorphisms in the Walker A motif that creates this mutation. This lysine residue is essential for nucleotide binding along with the main chain NH atoms.[36] Figure 2.3B shows a partial alignment comparing human, yeast and *P. falciparum* Rad 51 across this region. The conserved lysine is changed to arginine in the dominant negative form of the enzyme shown by the blue box. We expressed the Pf Rad51^{K143R} cassette from an episomal plasmid transfected into the 3D7 parasite line, these parasites are noted as R51DN. Previous work demonstrated that overexpression of Pf Rad51^{K143R} in a mouse model exposed to a chemical mutagen made the parasites more susceptible to DNA damaging agents as measured by growth in the host.[18] This Rad51DN plasmid expressed the dominant negative which had not been previously characterized in *P. falciparum*. Measurements of total Rad51 transcript levels by Q-RT-PCR detected an increase of approximately 50% between the wild-type and dominant negative expressing parasites, suggesting that the dominant negative is expressed at a level ~50% of endogenous Rad51 expression. While this is not an overexpression, small amounts of mutant PfRad51 protein were found to substantially reduce the ATPase activity of wild-type PfRad51 protein.[18] Roy et al reported that at a ratio of 8:1 wild-type to mutant, ATPase activity was inhibited 46% and a ratio of 4:1 caused 84% inhibition.[37] The mutant HsRad51^{K133R} was also found to exhibit a dominant

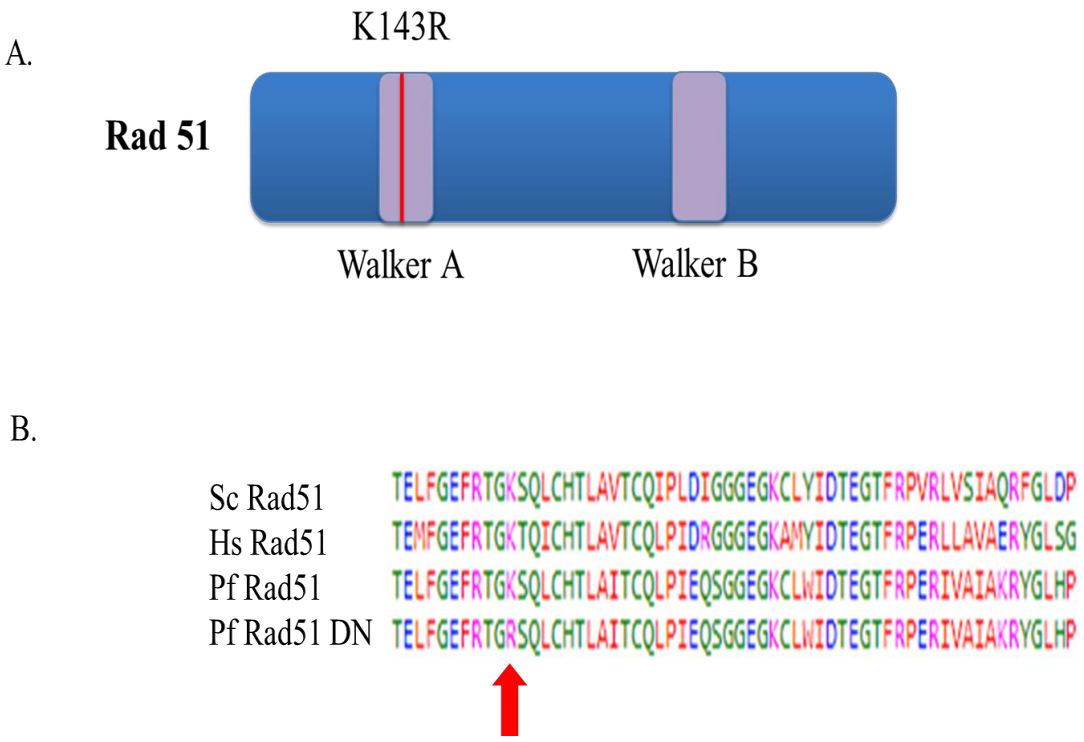


Figure 2.3.The Rad51 dominant negative mutation. A. Schematic of the Rad51 protein with the dominant negative mutation (K143R) in the Walker A motif (red line). B. A partial protein alignment of Rad51 with red arrow indicating the conserved lysine that is mutated in the Pf Rad51 DN. Sc Rad51- *Saccharomyces cerevisiae*, Hs Rad51- *Homo sapien*, Pf Rad51- *Plasmodium falciparum*

negative effect at low levels[20], indicating that low mutant expression can cause a dominant negative phenotype.

2.2.4 Increased Irradiation sensitivity of R51 DN parasites.

Given the absence of the C-NHEJ pathway, HR is the dominant repair pathway for DSBs in Plasmodium parasites. Rodent parasites harboring the dominant negative mutation failed to repair DSBs;[18] however similar work has not been performed with human malaria parasites. Therefore we decided to examine how the PfRad51^{K134R} dominant negative mutation affects *P. falciparum*'s ability to repair DSBs. We anticipated that the dominant negative Rad51 would interfere with HR and thus cause hypersensitivity to DNA damage. We irradiated R51DN parasites and used WT/DHOD parasites as control so that both parasite lines would contain a DHOD drug resistance cassette. The parasites were synchronized and seeded as described above.

R51DN trophozoites have higher irradiation sensitivity than WT/DHOD trophozoites and show the greatest difference in recovery to normal growth at 100 Gy irradiation. At 80Gy R51DN trophozoites recover to 2% parasitemia on average 2 days later then WT/DHOD, as R51DN average recovery is 6 days (+/- 2.61) while WT/DHOD is 4 days (+/- 2.34) (Figure 2.4B). The effect of the dominant negative is not significant at 80 Gy. At 100Gy though R51DN trophozoites require almost double the amount of time to recover normal growth as WT/DHOD indicating that an increase of 20Gy of irradiation does not mean a proportional change in time to normal growth recovery, as the recovery time is 14 days (+/- 2.66) for R51DN and 9 days (+/- 0.88)

for 100Gy WT/DHOD trophozoites. At 100 Gy irradiation the difference in recovery time between WT/DHOD and R51 DN is greater for trophozoites as compared to rings, trophozoites exhibiting a 5 day recovery difference while rings have a 3 day discrepancy. This is reflected in the P values for the two stages: 100 Gy rings is significant (< 0.05) and the difference between WT/DHOD and R51DN 100 Gy trophozoites is highly significant at < .001 This statistically significant difference in time to recovery is also represented in the growth curves where the slopes of the curves and days between curves on the plot corroborate the more attenuated growth phenotype of R51DN trophozoites as compared to R51DN rings. (Figure 2.4B).

For ring stage parasites a significant difference in irradiation sensitivity between WT/DHOD and R51DN was seen at 100 Gy exposure but not at 80 Gy (Figure 2.4A). After 80 Gy exposure, R51DN parasites at ring stage take 12 days (+/- 4.7) to recover from irradiation while WT/DHOD requires 9 days (+/- 2.39), statistical analysis indicates that the dominant negative phenotype is not significantly different. The difference in irradiation sensitivity is more apparent at 100Gy of irradiation, R51DN ring parasites needing 19 days (+/- 2.57) to return to normal growth as WT/DHOD 16 days (+/- 2.76) (Figure 2.4A). The difference in growth rates between WT/DHOD and R51DN at 100 Gy is statistically significant. The higher level of irradiation-induced DNA damage by 100 Gy is further exacerbated by the PfRad51^{K134R} mutation inhibiting the ability of the parasite to repair DSBs which is represented by the attenuated growth phenotype.

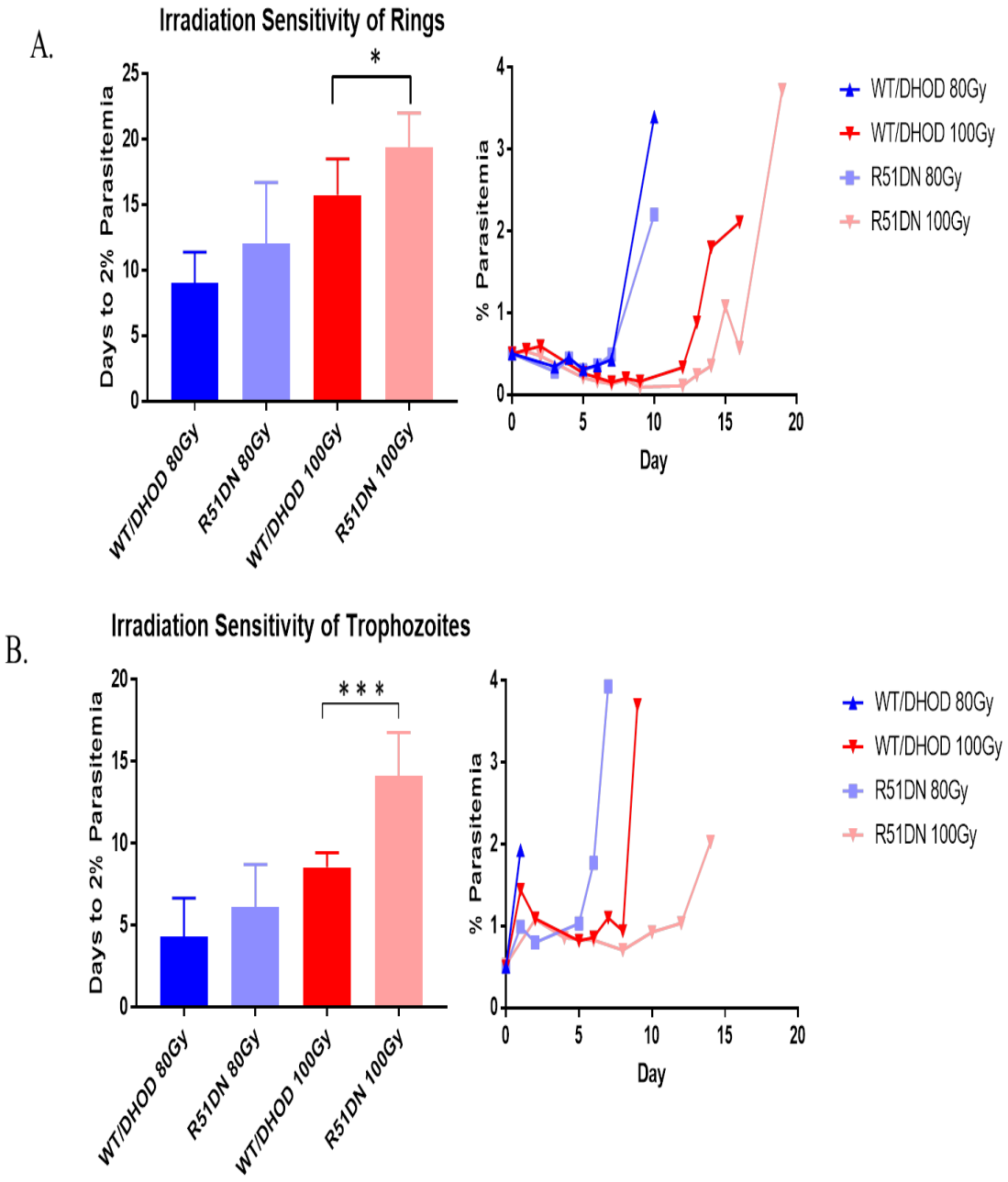


Figure 2.4. Increased Irradiation sensitivity of R51DN parasites A. Ring Irradiation sensitivity of WT/DHOD as compared to Rad51DN parasites. *- $P<0.05$, using two-way ANOVA with Sidak's comparison test. The growth curve is representative of one experiment with three replicates, three independent experiments done overall. **B.** Trophozoite Irradiation sensitivity of WT/DHOD as compared to Rad51DN parasites. ***- $P<0.001$ using two-way ANOVA with Sidak's comparison test. The growth curve is representative of one experiment with three replicates, three independent experiments were done overall.

2.3 Discussion

This work presents evidence that the ability of *P. falciparum* to repair DSBs is influenced by Rad51 function. The increased irradiation sensitivity of R51DN parasites indicates a necessary role for HR-mediated repair in response to irradiation DNA damage. The effects of the PfRad51^{K134R} dominant negative mutation were found to be stage and dosage dependent. Compared to control parasites, R51DN trophozoites showed a highly significant reduction in growth at 100Gy and a significant difference at 80 Gy. In comparison, rings only showed a significant change in recovery between dominant negative and wild type at 100 Gy, and this difference was less pronounced than what was observed for trophozoites. The increased dosage of irradiation, presumably leading to increased DNA damage, caused a more pronounced growth phenotype in both ring and trophozoite R51DN parasites. This reinforces the idea that HR is the dominant DSB repair pathway, necessary for repairing DSBs [18]. The toleration of irradiation by the trophozoite stage allows a clearer analysis of the R51DN phenotype while the higher irradiation sensitivity of the ring stage diminishes the impact the PfRad51^{K134R} dominant negative mutation has on the ring parasites. Our experiments have shown that DNA damage created by irradiation is repaired at different rates in a stage specific manner and repair is also dosage dependent in wild type parasites which corroborates previous findings.[26] The parasites are most irradiation sensitive at ring and fully segmented schizont stages while trophozoites are the least sensitive. *Plasmodium* species have asynchronous nuclear division[22] during schizogony which can make it hard to determine the timing between DNA synthesis and mitosis[38] to accurately identify different points in the cell

cycle. This allows for access to sister chromatids during trophozoite and early schizont phase, when DNA synthesis is at its peak[39], providing a homologous template for HR repair. The lack of homologous template during ring and segmented schizont stage begs the question of how these parasites repair DNA after significant DNA lesions caused by irradiation. The possibility that these parasites are performing A-NHEJ[5] at low rates could be tested by sequencing cloned parasites post irradiation and looking for the hallmarks of SD-MMEJ[40]. The mechanism by which these haploid parasites are able to repair DNA DSBs is not understood and research has yet been unable to identify an efficient alternative pathway to explain this parasite survival. Previous work has hypothesized that parasites that survive a high dose irradiation treatment may have been unaffected by the irradiation and sustained no damage.[41] It is possible that our time to recovery correlates with a small population of parasites regrowing undamaged but Southern and qPCR analysis of genomic DNA of similarly irradiated parasite cultures would indicate that a majority of parasites seen post-irradiation sustained DNA damage as evidenced by the inability to detect genes of interest and southern band deletions.(data not shown)

Irradiation can also cause cell cycle arrest, stalling the development of parasites which could mirror the regrowth of parasites with DNA damage. Ionizing irradiation can cause cell cycle delay usually in G_1 phase, before DNA replication or G_2 phase, before mitosis in eukaryotic cells.[42, 43] Cells then can reenter the cell cycle after repairing DNA damage, if present, stabilizing expression of cell cycle proteins and clearance of damaged proteins.[42] The possibility of parasite cell cycle arrest means the population of parasites that survive irradiation and the amount of days to

2% parasitemia could also be a cause of delayed growth due to cell cycle arrest and not a direct correlation of the parasite's ability to repair DNA damage. Determining which parasites had delayed regrowth due to cell cycle arrest or DNA damage is difficult as this would require single-cell analysis of a very small population of parasites (approx. ≤ 100 parasites). Cell cycle arrest in *Plasmodium* has been studied in response to artemisinin treatment[44] and artemisinin resistance[45], indicating that artemisinin-induced dormancy functions through a cell cycle arrest mechanism[44] and can allow for repair of artemisinin-induced DNA damage. *P. falciparum* artemisinin resistant isolates had a deceleration of blood stage development and this ring stage arrest may be associated with the parasite's ability to withstand drug pressure and regrow in favorable condisitons[46]. Cell cycle arrest can occure regardless of DNA damage but the mechanism is often utilized to allow parasite's to repair DNA before proceeding through the cell cycle. Though this evidence indicates that cell cycle arrest could account for the attenuated regrowth of parasites after irradiation it is unlikely that all surviving parasites were in cell cycle arrest and sustained no DNA damage. As such, irradiation is a widely accepeted method for creating DNA damage and studying repair.

P. falciparum is able to survive higher irradiation doses then *Toxoplasma gondii*[47], an Apicomplexan that has a robust NHEJ pathway. *T. gondii* oocysts dosed at 250 Gy of gamma irradiation were killed and unable to infect mice.[47] Our lab preliminarily irradiated *T. gondii* extracelleular tachyzoites and found their maximum survivable dose for reinvasion was about 30 Gy X-ray irradiation. (data not shown) While *T. gondii* tachyzoite Ku80 deletion mutants, inhibiting NHEJ activity had little

to no viability at 35 Gray gamma irradiation.[48] *P. falciparum* in comparison was found to survive 60 krad (600 Gy).[28] The higher irradiation tolerance of *P. falciparum* is peculiar as the parasite cannot perform NHEJ and at certain stages does not have the ability to perform HR while *T. gondii* has both NHEJ and HR pathways.[49]

Rad51 is an essential protein for repairing DSBs in most eukaryotes and in mammalian cells disrupting the function of Rad51 is lethal.[17] Deleting Rad51 in yeast is not lethal, but mutants are defective in repairing damage caused by irradiation and MMS, and in mitotic recombination.[50] There are no published reports of attempts to knockout PfRad51 therefore it is not clear whether the protein is necessary for viability in malaria parasites. The increased irradiation sensitivity of R51DN parasites indicates that the protein is a significant factor in HR repair, as expected. Previous work showed an inability of PfRad51^{K143R} rodent malaria parasites to recover from exposure to MMS, a chemical mutagen that also leads to DSBs, as opposed to the slower recovery of growth we observed in response to X-ray exposure.[18] A possible explanation is that low levels of endogenous WT PfRad51 are still present allowing for DSB repair, as the length of our experimental recovery period might allow for this. But in a previous study by Roy et al. the dominant negative protein was shown to inhibit the WT ATPase activity, showing inhibition of 46% with a ratio of 8:1 mutant to wild-type[18] which would likely be occurring in our *P. falciparum* model. Also HsRad51^{K133R} at low levels was found to give dominant negative phenotype and the 'rouge mixed filament model' of the protein would indicate that a mixed mutant and WT Rad51 nucleoprotein filament would be active but have improper Rad51 function.[20]

Another explanation is that our experimental design allows parasites to recover to normal growth over a several week period while Roy et al. examined DSB repair by checking for cell death after only several days. A longer time course for the *P. berghei* cultures could help elucidate whether Rad51 mediates DSB repair but is not essential. Creating a true PfRad51 knockout would definitively determine if HR in Plasmodium must be mediated by Rad51 and whether such a deletion is viable. The DSB repair pathway is missing proteins such as Rad52 [11] and RPA2 [1] which are canonical repair proteins in mammalian and yeast cells but the parasite seems to perform HR proficiently without these components. Finding homologs of missing canonical proteins may be difficult as they may are only be identifiable by structural domains or secondary structures and have low sequence similarity.[51, 52]

Rad51's role in HR repair becomes clinically relevant since gene conversion events caused by HR are what drive antigenic variation of the multi-copy gene families[53] which are essential virulence factors for the parasite. Understanding how to disrupt HR repair could have direct implications for developing an effective clinical vaccine. Our findings examining the Rad51 ^{K143R} dominant negative mutation in *P. falciparum* elucidated the irradiation sensitivity phenotype which indicates that HR is Rad51 mediated, but the protein may not be essential for repair. Creating a true Rad51 knockout will help to further explain Rad51's role in HR repair and provide a guideline for genetic targeting that could inhibit that repair. This will further therapeutics discovery to inhibit the parasite from repairing DNA and undergoing antigenic variation.

2.4 Materials and Methods

Parasite culture: *P. falciparum* parasites were cultivated at 5% hematocrit in RPMI 1640 medium (Corning Life Sciences, Tewksbury, MA), 0.5% albumax II (Invitrogen), 0.25% sodium bicarbonate, and 0.1 mg/ml gentamicin. Cultures were maintained at 37 °C in an atmosphere of 5% oxygen, 5% carbon dioxide, and 90% nitrogen. Individual clonal lines were obtained by limiting dilution.[54]

Creation of WT/DHOD and R51DN line. The 3D7 parasite line was transfected with a pINT plasmid containing either the DHOD cassette conferring resistance to DSM1 (Vaidya laboratory, Drexel University) as a control or a plasmid containing the dominant negative Rad51^{K143R} cassette[18]. The latter plasmid was created by inserting the HA tagged Rad51^{K143R} cassette which has a mutation in the Walker A motif which was synthesized by GeneWiz (Brooklyn, NY) into the pINT plasmid which has a 0.6-kb *Plasmodium chabaudi* dihydrofolate reductase-thymidylate synthase (PcDT) 5′ untranslated region (UTR)[55]. The original cassette conferring neomycin resistance in the pINT plasmid was replaced by DHOD. Parasites were transfected using "DNA loaded" red blood cells as previously described.[56] The pINT-DHOD plasmid control parasite line is referred to as WT/DHOD and the Rad51^{K143R} plasmid parasite line is referred to as R51DN. Both WT/DHOD and R51DN parasites were grown on 1.5 μM DSM1.

*Irradiation of parasite*s**.** All parasites were irradiated in a Rad Source 2000 irradiator set at 160 kV/25 mA following the manufacturer's specifications. Initial radiation tolerance curves were established with the 3D7 parasite line where it was determined

that a dose of 100 Gy led to significant growth impairment indicating widespread radiation damage, and that the parasite culture would return to normal growth.

WT/DHOD and R51DN parasite lines were synchronized with 5% (w/v) sorbitol which isolates ring stage parasites.[57] Parasites in blood pellet were incubated with 5% (w/v) sorbitol for 15 minutes at 37°C shaking. Parasites were rinsed in media twice and then reconstituted in culture.

Ring stage and trophozoite parasites at 0.5% parasitemia were irradiated with a Rad-Source Irradiator. Parasites were irradiated at 50, 80 and 100Gy and returned to standard culture conditions. Parasitemias counted by FACS was recorded until the cultures reached 2% infected RBCs.

Synchronization and irradiation of schizonts. All parasite lines were synchronized with 5% (w/v) sorbitol twice (as described above) followed by isolation of late-stage parasites by 40/70% Percoll-sorbitol gradient centrifugation[58]. For studies of late stage schizonts, parasite cultures were synchronized as above followed by treatment with a cysteine protease inhibitor, 10uM trans-Epoxysuccinyl-L-leucylamido(4-guanidino)butane (E-64 (Sigma- Aldrich, Brooklyn, NY) once they had reached a late trophozoite stage to prevent schizont rupture[31, 32]. Parasites were followed by microscopy to ensure they had all reached the segmented schizont stage and then they were irradiated at 100 Gy as above. Post irradiation, E-64 was removed with two washes with complete media. A 3D7 culture treated with E-64 but unirradiated was used as a control.

References

1. Lee, A.H., Symington, L.S. & Fidock, D.A. DNA repair mechanisms and their biological roles in the malaria parasite Plasmodium falciparum. *Microbiol Mol Biol Rev* **78**, 469-86 (2014).
2. Percario, S. et al. Oxidative stress in malaria. *Int J Mol Sci* **13**, 16346-72 (2012).
3. O'Neill, P.M., Barton, V.E. & Ward, S.A. The molecular mechanism of action of artemisinin--the debate continues. *Molecules* **15**, 1705-21 (2010).
4. Frankenberg-Schwager, M. & Frankenberg, D. DNA double-strand breaks: their repair and relationship to cell killing in yeast. *Int J Radiat Biol* **58**, 569-75 (1990).
5. Kirkman, L.A., Lawrence, E.A. & Deitsch, K.W. Malaria parasites utilize both homologous recombination and alternative end joining pathways to maintain genome integrity. *Nucleic Acids Res* **42**, 370-9 (2014).
6. Singer, M. et al. Zinc finger nuclease-based double-strand breaks attenuate malaria parasites and reveal rare microhomology-mediated end joining. *Genome Biol* **16**, 249 (2015).
7. Symington, L.S. Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol Mol Biol Rev* **66**, 630-70, table of contents (2002).
8. Nimonkar, A.V. et al. BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev* **25**, 350-62 (2011).
9. Sugiyama, T., Zaitseva, E.M. & Kowalczykowski, S.C. A single-stranded DNA-binding protein is needed for efficient presynaptic complex formation by the Saccharomyces cerevisiae Rad51 protein. *J Biol Chem* **272**, 7940-5 (1997).
10. Bhattacharyya, M.K. & Kumar, N. Identification and molecular characterisation of DNA damaging agent induced expression of Plasmodium falciparum recombination protein PfRad51. *Int J Parasitol* **33**, 1385-92 (2003).
11. Bhattacharyya, M.K., Bhattacharyya nee Deb, S., Jayabalasingham, B. & Kumar, N. Characterization of kinetics of DNA strand-exchange and ATP hydrolysis activities of recombinant PfRad51, a Plasmodium falciparum recombinase. *Mol Biochem Parasitol* **139**, 33-9 (2005).
12. Voss, T.S., Mini, T., Jenoe, P. & Beck, H.P. Plasmodium falciparum possesses a cell cycle-regulated short type replication protein A large subunit encoded by an unusual transcript. *J Biol Chem* **277**, 17493-501 (2002).
13. Moore, C.W. Responses of radiation-sensitive mutants of Saccharomyces cerevisiae to lethal effects of bleomycin. *Mutat Res* **51**, 165-80 (1978).
14. Vishnevetskaia, O., Luchnik, A.N., Arutiunova, L.S. & Shestakov, S.V. [Disorder of the repair of DNA double-stranded breaks in radiosensitive mutants of Saccharomyces cerevisiae yeasts]. *Genetika* **19**, 26-32 (1983).
15. Jachymczyk, W.J., von Borstel, R.C., Mowat, M.R. & Hastings, P.J. Repair of interstrand cross-links in DNA of Saccharomyces cerevisiae requires two systems for DNA repair: the RAD3 system and the RAD51 system. *Mol Gen Genet* **182**, 196-205 (1981).
16. Tsuzuki, T. et al. Targeted disruption of the Rad51 gene leads to lethality in embryonic mice. *Proc Natl Acad Sci U S A* **93**, 6236-40 (1996).
17. Lim, D.S. & Hasty, P. A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. *Mol Cell Biol* **16**, 7133-43 (1996).

18. Roy, N. et al. Dominant negative mutant of Plasmodium Rad51 causes reduced parasite burden in host by abrogating DNA double-strand break repair. *Mol Microbiol* **94**, 353-66 (2014).
19. Stark, J.M. et al. ATP hydrolysis by mammalian RAD51 has a key role during homology-directed DNA repair. *J Biol Chem* **277**, 20185-94 (2002).
20. Kim, T.M. et al. RAD51 mutants cause replication defects and chromosomal instability. *Mol Cell Biol* **32**, 3663-80 (2012).
21. Chi, P., Van Komen, S., Sehorn, M.G., Sigurdsson, S. & Sung, P. Roles of ATP binding and ATP hydrolysis in human Rad51 recombinase function. *DNA Repair (Amst)* **5**, 381-91 (2006).
22. Gerald, N., Mahajan, B. & Kumar, S. Mitosis in the human malaria parasite Plasmodium falciparum. *Eukaryot Cell* **10**, 474-82 (2011).
23. Lomax, M.E., Folkes, L.K. & O'Neill, P. Biological consequences of radiation-induced DNA damage: relevance to radiotherapy. *Clin Oncol (R Coll Radiol)* **25**, 578-85 (2013).
24. Cadet, J., Douki, T. & Ravanat, J.L. Oxidatively generated damage to the guanine moiety of DNA: mechanistic aspects and formation in cells. *Acc Chem Res* **41**, 1075-83 (2008).
25. Kondo, N., Takahashi, A., Ono, K. & Ohnishi, T. DNA damage induced by alkylating agents and repair pathways. *J Nucleic Acids* **2010**, 543531 (2010).
26. Waki, S., Yonome, I. & Suzuki, M. Plasmodium falciparum: attenuation by irradiation. *Exp Parasitol* **56**, 339-45 (1983).
27. Waki, S., Yonome, I. & Suzuki, M. X-ray sensitivity and DNA synthesis in synchronous culture of Plasmodium falciparum. *Z Parasitenkd* **71**, 213-8 (1985).
28. Oakley, M.S. et al. Radiation-induced cellular and molecular alterations in asexual intraerythrocytic Plasmodium falciparum. *J Infect Dis* **207**, 164-74 (2013).
29. Bennett, C.B. et al. Genes required for ionizing radiation resistance in yeast. *Nat Genet* **29**, 426-34 (2001).
30. Rothkamm, K., Kruger, I., Thompson, L.H. & Lobrich, M. Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol Cell Biol* **23**, 5706-15 (2003).
31. Chandramohanadas, R. et al. Apicomplexan parasites co-opt host calpains to facilitate their escape from infected cells. *Science* **324**, 794-7 (2009).
32. Glushakova, S., Mazar, J., Hohmann-Marriott, M.F., Hama, E. & Zimmerberg, J. Irreversible effect of cysteine protease inhibitors on the release of malaria parasites from infected erythrocytes. *Cell Microbiol* **11**, 95-105 (2009).
33. Absalon, S., Robbins, J.A. & Dvorin, J.D. An essential malaria protein defines the architecture of blood-stage and transmission-stage parasites. *Nat Commun* **7**, 11449 (2016).
34. Murray, L. et al. Multiplication rate variation in the human malaria parasite Plasmodium falciparum. *Sci Rep* **7**, 6436 (2017).
35. Sung, P. & Robberson, D.L. DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. *Cell* **82**, 453-61 (1995).
36. Hanson, P.I. & Whiteheart, S.W. AAA+ proteins: have engine, will work. *Nat Rev Mol Cell Biol* **6**, 519-29 (2005).
37. Roy, N. et al. Dominant negative mutant of Plasmodium Rad51 causes reduced parasite burden in host by abrogating DNA double-strand break repair. *Mol.Microbiol.* **94**, 353-366 (2014).
38. Naughton, J.A. & Bell, A. Studies on cell-cycle synchronization in the asexual erythrocytic stages of Plasmodium falciparum. *Parasitology* **134**, 331-7 (2007).

39. Inselburg, J. & Banyal, H.S. Synthesis of DNA during the asexual cycle of Plasmodium falciparum in culture. *Mol Biochem Parasitol* **10**, 79-87 (1984).
40. Yu, A.M. & McVey, M. Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res* **38**, 5706-17 (2010).
41. Adachi, K. et al. Postoperative administration of dienogest for suppressing recurrence of disease and relieving pain in subjects with ovarian endometriomas. *Gynecol Endocrinol* **32**, 646-649 (2016).
42. Bernhard, E.J., Maity, A., Muschel, R.J. & McKenna, W.G. Effects of ionizing radiation on cell cycle progression. A review. *Radiat Environ Biophys* **34**, 79-83 (1995).
43. Dimitrijevic-Bussod, M., Balzaretti-Maggi, V.S. & Gadbois, D.M. Extracellular matrix and radiation G1 cell cycle arrest in human fibroblasts. *Cancer Res* **59**, 4843-7 (1999).
44. Gray, K.A. et al. Correlation between Cyclin Dependent Kinases and Artemisinin-Induced Dormancy in Plasmodium falciparum In Vitro. *PLoS One* **11**, e0157906 (2016).
45. Mok, S. et al. Drug resistance. Population transcriptomics of human malaria parasites reveals the mechanism of artemisinin resistance. *Science* **347**, 431-5 (2015).
46. Witkowski, B. et al. Novel phenotypic assays for the detection of artemisinin-resistant Plasmodium falciparum malaria in Cambodia: in-vitro and ex-vivo drug-response studies. *Lancet Infect Dis* **13**, 1043-9 (2013).
47. Dubey, J.P., Jenkins, M.C., Thayer, D.W., Kwok, O.C. & Shen, S.K. Killing of Toxoplasma gondii oocysts by irradiation and protective immunity induced by vaccination with irradiated oocysts. *J Parasitol* **82**, 724-7 (1996).
48. Fox, B.A., Ristuccia, J.G., Gigley, J.P. & Bzik, D.J. Efficient gene replacements in Toxoplasma gondii strains deficient for nonhomologous end joining. *Eukaryot Cell* **8**, 520-9 (2009).
49. Kirkman, L.A. & Deitsch, K.W. Recombination and Diversification of the Variant Antigen Encoding Genes in the Malaria Parasite Plasmodium falciparum. *Microbiol Spectr* **2** (2014).
50. Paques, F. & Haber, J.E. Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiol Mol Biol Rev* **63**, 349-404 (1999).
51. Gopalakrishnan, A.M. & Kumar, N. Opposing roles for two molecular forms of replication protein A in Rad51-Rad54-mediated DNA recombination in Plasmodium falciparum. *MBio* **4**, e00252-13 (2013).
52. Badugu, S.B. et al. Identification of Plasmodium falciparum DNA Repair Protein Mre11 with an Evolutionarily Conserved Nuclease Function. *PLoS One* **10**, e0125358 (2015).
53. Claessens, A. et al. Generation of antigenic diversity in Plasmodium falciparum by structured rearrangement of Var genes during mitosis. *PLoS Genet* **10**, e1004812 (2014).
54. Kirkman, L.A., Su, X.Z. & Wellems, T.E. Plasmodium falciparum: isolation of large numbers of parasite clones from infected blood samples. *Exp Parasitol* **83**, 147-9 (1996).
55. Nkrumah, L.J. et al. Efficient site-specific integration in Plasmodium falciparum chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat Methods* **3**, 615-21 (2006).

56. Deitsch, K., Driskill, C. & Wellems, T. Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes. *Nucleic Acids Res* **29**, 850-3 (2001).
57. Lambros, C. & Vanderberg, J.P. Synchronization of Plasmodium falciparum erythrocytic stages in culture. *J Parasitol* **65**, 418-20 (1979).
58. Fernandez, V., Treutiger, C.J., Nash, G.B. & Wahlgren, M. Multiple adhesive phenotypes linked to rosetting binding of erythrocytes in Plasmodium falciparum malaria. *Infect Immun* **66**, 2969-75 (1998).

Chapter 3: Chromosome End Repair and Genome Stability

3.1 Introduction

Malaria, a mosquito transmitted infectious disease, continues to pose a major challenge for health care systems in many parts of the developing world [1]. The disease is caused by various species of the *Plasmodium* genus of protozoan parasites, members of the Apicomplexa family of obligate parasites that also includes *Toxoplasma*, *Babesia*, *Cryptosporidium* and *Theileria*. The pathogenic stage of malaria occurs when parasites invade and replicate within the red blood cells (RBCs) of their vertebrate hosts causing cycles of fever and chills as well as severe anemia. *P. falciparum*, the species responsible for the most virulent form of the disease, also causes infected RBCs to become cytoadhesive, resulting in sequestration within the post-capillary vasculature, leading to circulatory disruption and many of the severe symptoms of the disease including cerebral malaria and complications during pregnancy [2]. While multiplying within the RBCs the parasites are haploid, replicating through a process called schizogony in which multiple rounds of nuclear division lead to a multinucleated cell containing 16-32 individual merozoites. These are released upon lysis of the host cell enabling them to invade new RBCs and reinitiate the cycle. Throughout this process parasites are confronted with numerous conditions that can cause DNA damage, including the products of metabolism and hemoglobin digestion, components of the host immune response and oxidative damage resulting from exposure to antimalarial drugs [3-5]. The success of these parasites therefore depends on their ability to efficiently repair DNA damage and maintain genome integrity.

The *P. falciparum* genome is arranged into fourteen linear chromosomes constituting a total of approximately 23 Mb [6]. The contents of each individual chromosome are organized similarly, with single copy genes responsible for replication and progression through the cell cycle found within the central regions and members of the large, multi-copy, clonally variant gene families found in lengthy arrays within the subtelomeric domains. Flanking the variant genes are conserved telomere associated repeat elements (TAREs) which typically extend for 10-30 kb followed by the actual chromosome end consisting of typical telomeric heptad repeats maintained by telomerase (Figure 3.1A) [7]. The subtelomeric domains of *P. falciparum* are of particular interest for understanding parasite biology since the clonally variant gene families that reside there are the primary virulence determinants and their varied expression results in antigenic variation, the process that enables parasites to perpetuate long-term, chronic infections. These regions contain hundreds of genes from several gene families such as *var stevor*, *rifin*, *Pfmc-2TM*, *FIKK* and *acs*, which are maintained in a unique chromatin structure that is marked by the histone modifications H3K9me3 and H3K36me3 [8-10]. Within the nucleus these regions cluster at the nuclear periphery, an arrangement that is thought to facilitate recombination between members of the multi-copy gene families that reside on different chromosomes, thereby generating diversity [11, 12]. Given the importance of these chromosomal regions for pathogenesis, understanding how they are replicated, maintained and repaired is of considerable interest.

DNA double strand breaks (DSBs) are an especially severe form of DNA damage that must be repaired for a cell to remain viable and to replicate. Eukaryotic

cells typically rely on two distinct pathways to repair such breaks, homologous recombination (HR) and non-homologous end joining (NHEJ). While NHEJ can ligate the broken ends of the DNA back together to efficiently repair a break, it often results in small deletions at the site of ligation and can therefore be mutagenic [13]. HR is a more accurate method of repair since it copies information from homologous sequences elsewhere in the genome, typically the other allele in a diploid cell, thereby avoiding any insertion or deletion at the site of the break. HR however requires a homologous region of high sequence identity (typically greater than 98%) to serve as template for repair [14]. The asexual stages of *P. falciparum* represent a unique organism for the study of DSB repair since they are haploid and thus generally lack homologous sequences to serve as templates for HR and yet they are missing the canonical NHEJ pathway. An alternative NHEJ pathway based on microhomology has been described, although this appears to be quite inefficient [15, 16]. Therefore how malaria parasites respond to DSBs and efficiently maintain genome integrity is unclear. Characterization of important enzymes involved in HR demonstrate that this pathway is functionally conserved [17-19]. The subtelomeric regions, given their size, extensive gene content, semi-repetitive nature and unique chromatin structure, represent a particularly interesting region of the genome for studying DSB repair.

In addition to HR and NHEJ, eukaryotic organisms can also stabilize DSBs occurring at the chromosome ends through the action of telomerase, the enzyme that maintains the telomeric repeats at this specialized chromosomal location. This repair pathway, called de novo telomere addition or “telomere healing”, involves the recruitment of the telomerase complex directly to the site of the break [20, 21]. The DNA

strand is resected until a region of high TG content is encountered, which is thought to "seed" telomerase and enables it to incorporate telomeric repeats at the site of DNA break and resection (Figure 3.1B). This results in the creation of a functional telomere, thereby stabilizing the chromosome and maintaining genome integrity, albeit with a deletion of the region of the chromosome between the DSB and the original telomere and loss of the intervening genetic information. Telomere healing is thought to compete with HR and NHEJ for repair of DSBs, with the pathway of repair depending on the chromosomal environment in which the break occurs [20]. Telomere healing has been described for a number of eukaryotic organisms and the addition of telomeric repeats to the ends of broken chromosomes has been described for *P. falciparum* [22-24], indicating that this mechanism of repair is conserved. However, the details of telomere healing in malaria parasites and how this pathway integrates with HR and alternative NHEJ has not been studied extensively.

Given the general chromosome structure of *P. falciparum* (Figure 3.1A), telomere healing could repair DSBs that occur anywhere within the extensive subtelomeric regions between the chromosome ends and the internal, highly conserved regions of the genome. Genes within these regions are necessary for host-pathogen interactions but are not required for viability in cultured parasites, thus deletions of subtelomeric regions are tolerated and can be efficiently recovered and analyzed. The semi-repetitive nature of these regions suggests that DSBs could also potentially be repaired by HR, allowing the study of both pathways as they repair breaks within this chromosomal environment. Detailed sequence examination of subtelomeric domains has been difficult in the past due to the repetitive nature of the sequences, making

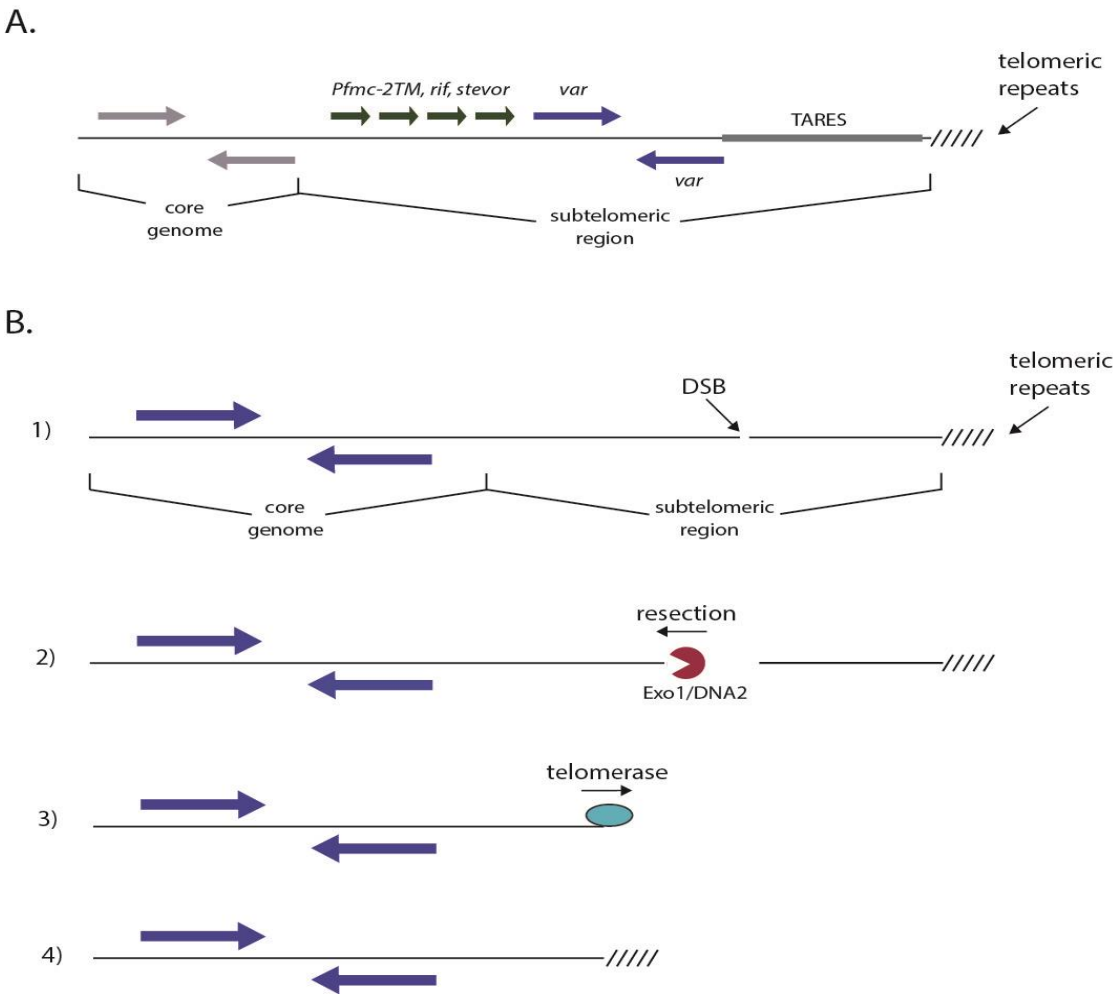


Figure 3.1. Repair of double strand breaks within subtelomeric chromosomal regions. A. The typical structure of the subtelomeric regions of the chromosomes of *P. falciparum*. The core genome contains primarily single copy housekeeping genes (gray) while the subtelomeric regions consists of large arrays of variant antigen encoding genes of the *var* (blue), *rif*, *stevor* and *Pfmc-2TM* (green) families. Telomere associated repeat elements (TARES) are positioned between the variant gene families and the telomere repeats. B. The steps of de novo telomere addition, also called telomere healing. 1) Typical chromosomes can be divided into a core genome containing housekeeping genes, telomere repeats at the extreme end of the chromosome, and intervening subtelomeric regions. A double strand break (DSB) within the subtelomeric region can be repaired by telomere healing. 2) The DSB is recognized by protein complexes that include exonuclease activity. In model organisms, both Exo1 and DNA2 have been implicated in the resection of DNA away from the telomere, revealing a single stranded 3' end. 3) When a single strand sequence is revealed that can anneal to the template RNA of the telomerase complex, telomerase activity extends from the break, placing telomere repeats directly at this site. 4) Repeated rounds of telomere addition results in a stable telomere and maintains genome integrity.

assemblies of these regions from short sequence reads problematic. However newer technologies including those that utilize single molecule real time (SMRT) sequencing enable the confident assembly of complete subtelomeric regions, thus allowing us to examine in detail DSB repair in these regions of the genome. We applied this technology to study the repair of spontaneous DSBs that occur within subtelomeric regions during in vitro culture as well as breaks induced randomly through exposure to ionizing radiation.

Our analysis indicates that both telomere healing and gene conversion through HR can repair DSBs within subtelomeric domains. Telomere healing was the most common type of repair we observed, and which pathway was utilized was strictly determined by the sequence surrounding the breakpoint. Considering the repetitive nature of these chromosomal regions, subtelomeric deletions resulting from telomere healing events could be followed by subsequent gene conversion through HR, and thus reestablish the typical subtelomeric structure and maintain the overall parasite chromosome organization. We conclude that both repair pathways therefore work in tandem to maintain genome integrity and preserve the complement of clonally variant genes found within the genomes of *P. falciparum* isolates.

3.2 Results

3.2.1 Extensive telomere healing observed in cultured parasites

While telomere healing has been described in *P. falciparum* [22-24], how commonly it occurs within the 28 subtelomeric domains of the parasite's genome has not been closely examined. Telomere healing has been most extensively studied in

yeast, where telomerase displays a strong preference for TG-rich sequences to initiate synthesis of repeats [25, 26]. If *P. falciparum* telomerase has a similar preference for TG-rich sequences to initiate synthesis of repeats, then one is likely to observe telomeric repeats fused to the coding regions of truncated genes at the new chromosome ends, as coding regions are significantly enriched in Cs and Gs compared to non-coding regions of the genome.

As a first step to determine if telomere healing is common in cultured *P. falciparum* and to obtain a baseline for our experimental analysis, we analyzed the sequence of all 28 chromosome ends provided in the reference genome sequence of 3D7 (Plasmodb.org). Five chromosome ends display the hallmarks of telomere healing, including both the absence of TAREs as well as telomeric repeat sequences fused directly into or next to protein coding regions. In four of these, on chromosomes 5, 6, and both ends of 14, the telomeric repeats are found within the coding region of a truncated *var* gene, creating a pseudogene. In the fifth example on chromosome 11, the telomeric repeats are found immediately downstream of an intact *var* gene (Figure 3.2A).

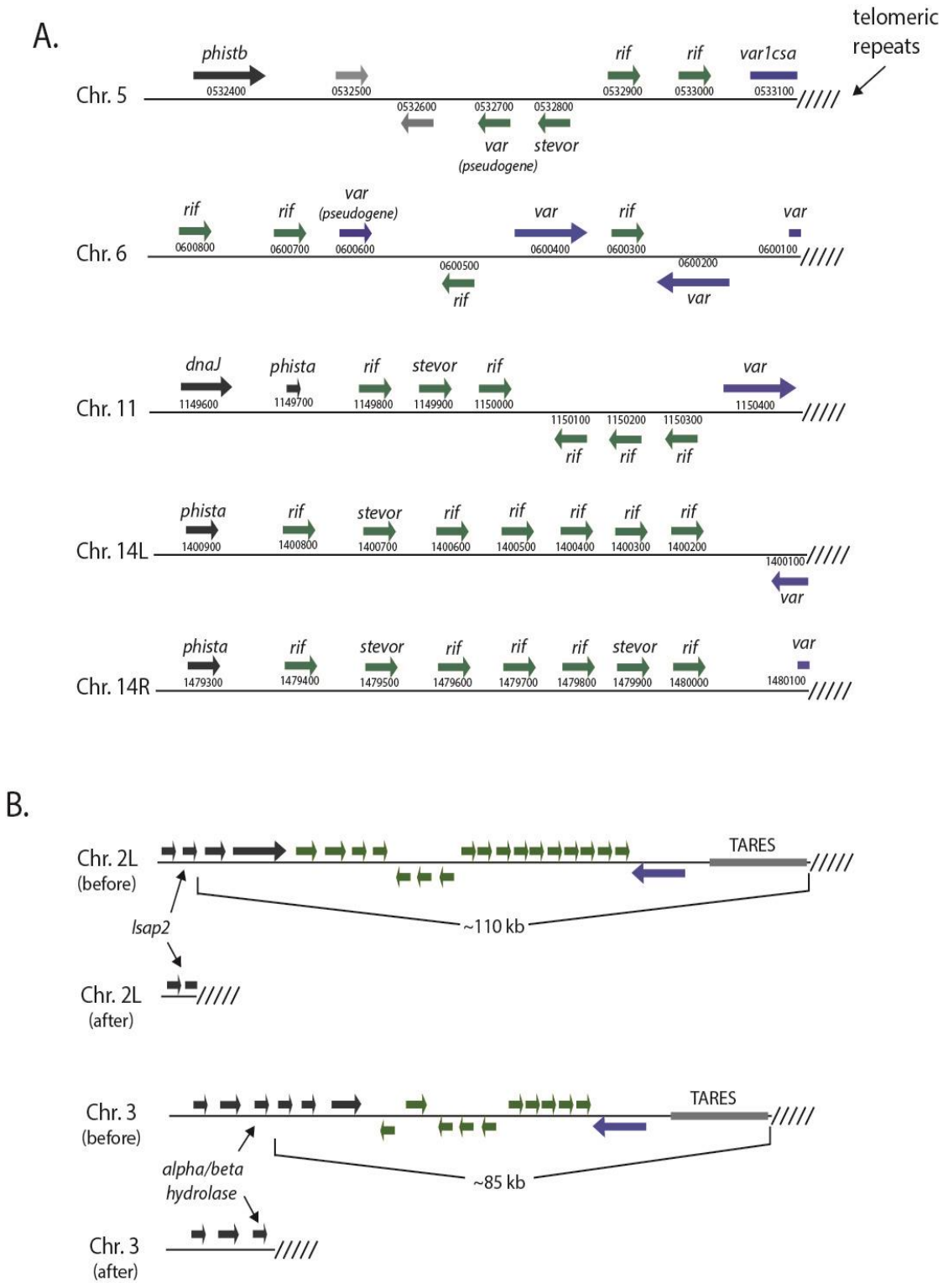
Next we investigated if telomere healing events occurred in parasites reared in the laboratory, we generated several sub-clones from a stock of 3D7 that had been grown in culture for several years in our laboratory and chose one for whole genome sequencing. To determine the sequence of the chromosome ends and to avoid difficulties in assembling the repetitive structure of subtelomeric domains from the short sequence reads typically derived from most methods of high throughput

sequencing, we instead utilized single molecule real time (SMRT) sequencing which can yield single contiguous reads of 10-20 kb [27, 28]. Using this method, we verified the five telomere healing events recognized in the reference genome sequence and also identified two additional novel DNA sequences consistent with telomere healing (Figure 3.2B and Figures 3.3, 3.4, 3.5). The first on chromosome 2 resulted in the deletion of approximately 110 kb, including the entire subtelomeric domain and leading to the insertion of telomeric repeats within the *lsap2* gene (PF3D7_0202100). The second occurred on chromosome 3 and similarly deleted all of the subtelomeric domain (~85 kb), inserting telomeric repeats just downstream of a gene encoding a putative alpha/beta hydrolase (PF3D7_0301300). These two events provide us with a "before and after" picture of telomere healing, and further indicate that this is a common mechanism of DSB repair within the subtelomeric regions of *P. falciparum*.

3.2.2 Inducement of DSBs and repair in cultured parasites by exposure to X-ray irradiation

To more directly observe DSB repair within subtelomeric domains, we chose to induce random DSBs by exposing parasite cultures to X-ray irradiation. X-rays are known to cause DSBs without sequence bias, and such breaks must be repaired for parasite viability. Thus, by selecting for viable parasites after near lethal exposure to

Figure 3.2. Telomere healing events identified in the genome of the parasite 3D7. A. The five telomere healing events identified in the genome of the reference 3D7 sequence. The chromosome number is shown along with the gene structure of each subtelomeric region. For each gene, the gene family is indicated as well as the annotation number as provided by PlasmoDB.org. B. Two additional telomere healing events identified in a clonal line of 3D7. The event that occurred on chromosome 2 (top) results in deletion of ~110 kb and leads to truncation of the *lsap2* gene (PF3D7_0202100) while the healing event on chromosome 3 resulted in loss of ~85 kb and added telomere repeats just downstream of a gene encoding a predicted alpha/beta hydrolase (PF3D7_0301300).


A.
telomeric repeats
Chr. 5
phistb
0532400
0532500
0532600
0532700
0532800
var (pseudogene)
stevor
rif
0532900
rif
0533000
var1csa
0533100
Chr. 6
rif
0600800
rif
0600700
var (pseudogene)
0600600
0600500
rif
var
0600400
rif
0600300
0600200
var
var
0600100
Chr. 11
dnaJ
1149600
phista
1149700
rif
1149800
stevor
1149900
rif
1150000
1150100
rif
1150200
rif
1150300
rif
var
1150400
Chr. 14L
phista
1400900
rif
1400800
stevor
1400700
rif
1400600
rif
1400500
rif
1400400
rif
1400300
rif
1400200
1400100
var
Chr. 14R
phista
1479300
rif
1479400
stevor
1479500
rif
1479600
rif
1479700
rif
1479800
stevor
1479900
rif
1480000
var
1480100
B.
Chr. 2L (before)
TARES
lsap2
~110 kb
Chr. 2L (after)
Chr. 3 (before)
TARES
alpha/beta hydrolase
~85 kb
Chr. 3 (after)

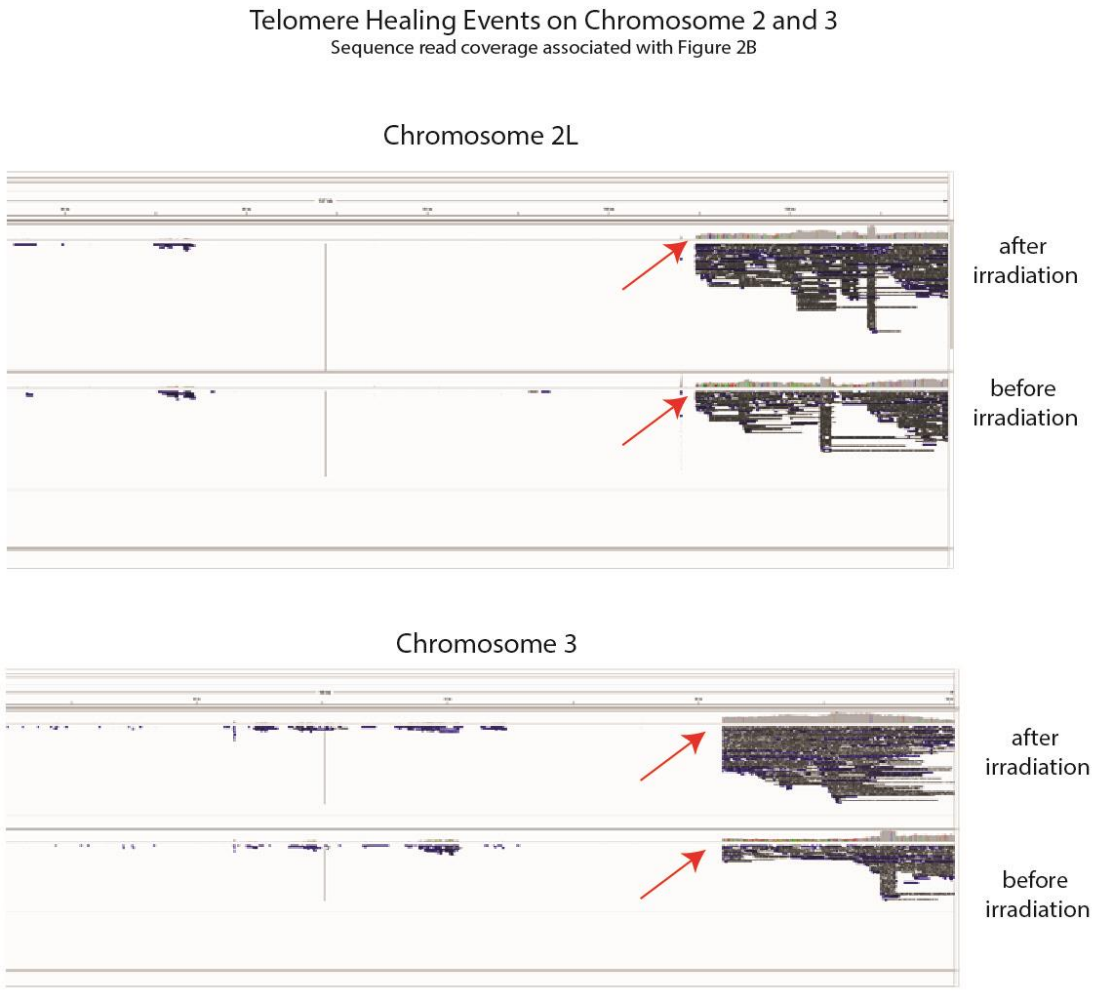


Figure 3.3. Sequence read "pile-ups" displaying the detection of telomere healing events on chromosomes 2 and 3. These events are displayed schematically in Fig. 3.2B of the main text. For both chromosomes, the reads obtained from the irradiated clone are shown on the top, and the reads obtained for the nonirradiated clone are shown on the bottom. All reads were aligned with the 3D7 reference sequence obtained from Plasmodb.org. The sites of telomere healing are denoted with a red arrow. The subtelomeric deletions were detectable in both the irradiated and nonirradiated clones, indicating that these two healing events occurred during culture of the parent line prior to exposure to X-ray irradiation. Reads that aligned within the deleted regions in the irradiated line coincide with TAREs that are shared between many subtelomeric domains.

ATGAAGGGATCTGGATCAGAAAAAAATGTATATCTTTCAAATAAAAATAAAGAAATTAATATGAACCAACAATCAG
ATAATAAAATGTGTGATGAATGTGATGATATGAATCAACCAGGAGATGTAAATAAAAATGACAAAACATCAAATG
ATCAAGCAAATTCAAGTGATTCTGATTGTGAGCCCTTACCATTTGGATTAAAACCTTCAGATTTAAATAGAAAAGTT
ACAGAAGAAGATTTAGAAAGAATGATAATAGAATTACCAGGAAAATTAGAAAGGAAAGATATGTATTTAATATGG
CATTATAGTCATTCTCTTTTGAGAGATAAATTTAATAAAATGAAAAGTTCGTTATGGAGTATTTGTGGGAAATTAGC
TCATGAACATAAGTTACCATTCAAAATTAAAATGAAGAAATGGTGGAAATGTTGTGGTCATGTTACAGATGAATTA
TTAATAAAAGAGCATGATGATTATAATTCTATATATAATTATATTAATAATGAATCATCAAGTCGTGAACAATTTCTT
ATATTTCTTAATATGATAAAGCATTCATGGACAACATTTACTATGGAGACTTTTATTAAATGTAAGATTTCTTTAGAA
AATAACATGAGAAATGTTACAGGGTTTAGGGTTCAGGGTTTAGGGTTTAGGGTTCAGGGTTTAGGGTTTAGGGTT
TAGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGGTTCAGGGTTTAGGTTTAGGGTTTAGGGTTCAGGGT
TTAGGTTTAGGGTGTAGGGTTCAGGGTTTAGGGTTTAGGGTTCAGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGG
TTTAGGGTTCAGGGTTTAGGGTTCAGGGTTTAGGGTTCCAGGGTTCAGGGTTCAGGGTTTAGGGTTCAGGGTTTA
GGGTTCAGGGTTTTAGGGTTCAGGGTTCAGGGTTTAGGGTTTAGGGTTCAGGGTTTAGGGTTCAGGGTTCAGGGT
TTAGGGTTTAGGGTTCAGGGTTTAGGGTTCAGGGTTCAGGGTTTAGGGTTCAGGGTTTAGGTTTTAGGTTTAGGG
TTCAGGGTTTAGGGTTTTAGGTTTAGGGTTCAGGGTTTAGGGTTCAGGGTTTAGGGTTTAGGGTTCAGGGTTCAG
GGTTTAGGGTTTAGGGTTCAGGGTTTAGGGTTCAGGGTTCAGGGTTTAGGGTTCAGGGTTTAGGGTTCAGGGTTC
AGGGTTTAGGGTTCAGGGTTTAGGGTTCAGGGTTCAGGTTTAGGGTTTAGGGTTTAGGGTTTAGG

Figure 3.4. Assembled sequence showing the telomere healing event associated with the end of chromosome 2L. Telomere healing event shown schematically in Fig. 3.2B of the main text. The coding region of the gene lsap2 is shown in black text, while the telomeric repeats are shown in blue.

CCCTAAACCCTGAACCCTAAACCCTAAACCCTAAACCCTAAAACCCTGAACCCTAAACCTAAACCCTGAACCCTAAA
CCCTGAACCCTAAACCTAAACCCTAAACCCTAAACCCTGAACCTAAACCTAAACCCTGAACCCTAAACCCTAAACCC
TAAACCTAAACCCTAAACCCTAAACCCTGAAACCCTGAACCCTGAACCCTAAACCCTAAACCCTAAACCCTAAACCC
TGAACCCTAAACCTAAACCCTAAACCCTAAACCCTAAACCCTAAACCCTAAACCCTAAACCCTGAACCCTAAACCCT
GAAACCCTAAACCCTAAACCCTAAACCCTAAACCCTAAACCCTGAAACCCTAAACCTAAACCCTAAACCCTAAACCC
TGAACCCTGAACCCTAAACCCTGAACCCTAAACCCTGAACCCTAAACCCTGAAAAAAAAAAAAAAAAAAAAAAAAGT
ATGTACAATGATAATAAATATATATATATATAAAAAAAAAAAAGGGTATAGACCACATAATATTATATTTTATTTTT
TTTTTTTCTTATATATTATGTGGTACATCCATTTTTATCAACAGGGAAATAAATATCATCTGATTTTAAATTCACATCA
TTTGGAAAAGTGGTAAAGATTGCAATACATAGAATAATATCTGAAGTTCTTGAAGGTATTATATAATGATTTTCTCC
TTTGAAAACAATAATATGAGAATTAGTAAAATAGGTATTTAAATTTTTATAAACATCGTCATCATAAAGATCATCTT
TTTCTCCACAAAATATTAAAGTAGGTATATTTTTTTTCCCTATTTCTATAAAAATGTCATGGGAACTCCACATTGGCA
TATTATTTAAACAACCATATATACAATCGGTTAAATTCTTTTTGGCAAAAGCATTCCACATTAATTTATCATAAACTA
CTTCAAAATCATTTCGTGGAATACAATGTTTCATACAACATTTGGGAAAACAAAATGGTGACATAATACATGAACA
AGAATTAATTATACTTAATAATCCATTTCTTATTTTTAATGATAAAGGTTTCCTTCCTAACATACCCACAGGTGATAA
AAAAACAATTTTTTTAACTTGATTAATATATTTCTGTGCAAATGCAGCAGCTATTAAACAACCCATTGAACCACCAAT
CAAATAAAATTCCTTATTTTGCAAATTTAAATAACATACTAATTCTTCTATTTGAGTCAAAAAAAAATTTAAATTATA
TACATCATCACTACTATATTTTGGACATTCAGATAACCCATGTCCATATAAATCATACTTAAGTATTTGATAATTATT

Figure 3.5. Assembled sequence showing the telomere healing event associated with the end of chromosome 3. The telomere healing event shown schematically is in Fig. 3.2B of the main text. The coding region of the alpha/beta-hydrolase gene is shown in black text, while the telomeric repeats are shown in blue. A short stretch of subtelomeric DNA between the coding region and the telomere repeats is shown in purple text.

X-rays, we hoped to detect and analyze additional examples of DSBs within subtelomeric regions. The subclone of 3D7 previously used for SMRT sequencing was exposed to increasing amounts of X-ray irradiation to determine the level of exposure that would lead to significant widespread DNA damage yet allow for parasites to repair, recover and propagate. Exposure to 100 Gy resulted in significant lethality, however viable parasites outgrew from irradiated cultures within 10-12 days after exposure. Parasites were exposed to 100 Gy three times consecutively, allowing the parasites to recover normal growth after each irradiation. Subclones were then isolated by limiting dilution and one clone subjected to whole genome sequencing using the SMRT methodology. Analysis of the genome sequence identified two additional examples of telomere healing, one on chromosome 1 leading to a deletion of ~90 kb and resulting in the insertion of telomeric repeats into a *rifin* gene (Pf3D7_0101900) and a second on chromosome 2 leading to deletion of ~100 kb and resulting in insertion of telomeric repeats just downstream of the hypothetical gene Pf3D7_0221000 (Figure 3.6A and Figures 3.7, 3.8 and 3.9). This clone of 3D7 now carries significant truncations within nine of its 28 subtelomeric regions.

In addition to the two new examples of telomere healing identified in this clone, we also detected three examples of recombination events likely resulting from HR. One of the subtelomeric regions of chromosome 13 is a hybrid sequence in which the original subtelomeric region has been deleted and replaced by ~25,500 bp, including the telomeric repeats, from one of the subtelomeric regions of chromosome 9. The breakpoint of the recombination event occurred within the coding region of a *var* gene, creating a new *var* gene that is a chimera of PF3D7_0900100 and

PF3D7_1300100 (Figure 3.6B). A more complex product of recombination was identified near one end of chromosome 12. The first ~8100 bp, including the telomeric repeats, are derived from one of the subtelomeric domains of chromosome 9 and, given that both ends of chromosome 9 remain unchanged, this appears to be the result of a gene conversion event. This fragment is fused to 13,886 bp of sequence identical to a region within one of the subtelomeric domains of chromosome 1. This sequence is within the portion of chromosome 1 that is now deleted in this clone, indicating that it was transposed into the subtelomeric region of chromosome 12 prior to its deletion from chromosome 1 (Figure 3.6B). This could have resulted from either gene conversion or reciprocal recombination. These data indicate DSBs that occur within subtelomeric domains can be repaired either by HR or telomeric healing, and that the repair process can generate new *var* genes.

3.2.3 Sequence preference for insertion of telomeric repeats by telomerase

A model for telomere healing in higher eukaryotic cells has been derived from extensive experiments conducted primarily in yeast. When a DSB forms within a subtelomeric region, the DNA is initially resected by one of two exonucleolytic pathways (Exo1 or Dna2/Sgs1) revealing a region of 3' single-stranded DNA [29]. The resection continues until a sequence is encountered that can anneal to the template region of telomerase RNA. This allows telomerase to begin synthesizing the telomeric repeats directly at the end of the chromosome, continuing for multiple rounds of DNA synthesis and resulting in a functional telomere that can stabilize the chromosome end

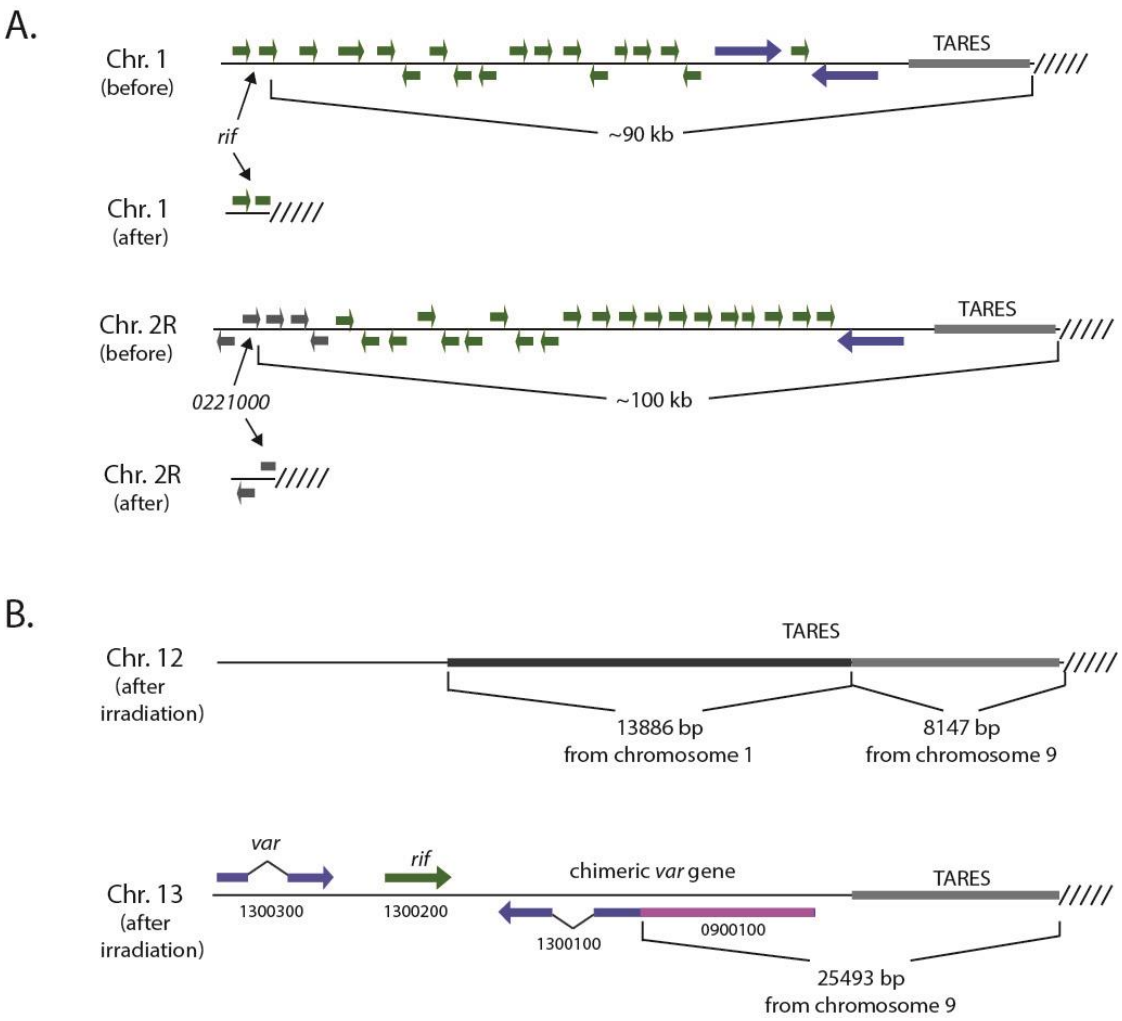


Figure 3.6. Double strand break repair within subtelomeric regions in response of exposure to X-ray irradiation. A. Two products of telomere healing are shown. On chromosome 1 (top), telomere repeats were inserted into the coding region of the *rif* gene Pf3D7_0101900 and on chromosome 2 (bottom) in which telomere repeats were inserted into the hypothetical gene Pf3D7_0221000. These events resulted in subtelomeric deletion of ~90 and ~100 kb, respectively. B. Three products of repair by homologous recombination were also identified, 2 on chromosome 12 (top) and one on chromosome (13). The two events on chromosome 12 resulted in the insertion of subtelomeric sequences from chromosomes 1 and 9, however no coding regions were altered. The event on chromosome 13 resulting in the inserted of 25,493 bp of sequence from chromosome 9 and lead to the creation of a new *var* gene that is a chimera of Pf3D7_1300100 (blue) and Pf3D7_0900100 (pink).

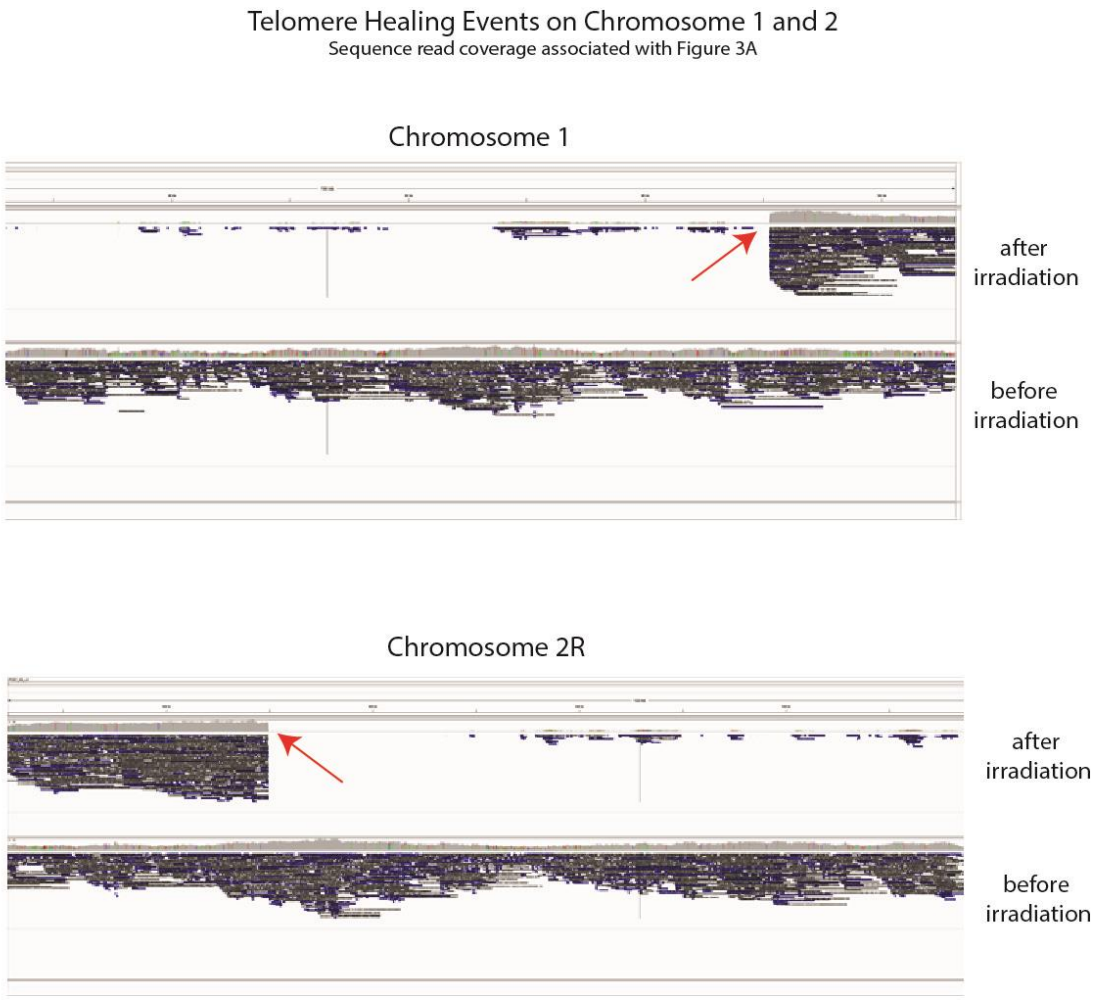


Figure 3.7. Sequence read "pile-ups," displaying the detection of telomere healing events on chromosomes 1 and 2. These events are displayed schematically in Fig. 3.6A of the main text. For both chromosomes, the reads obtained from the irradiated clone are shown on the top and the reads obtained from the nonirradiated clone are shown on the bottom. All reads were aligned with the 3D7 reference sequence obtained from Plasmodb.org. The sites of telomere healing in the irradiated clone are denoted with a red arrow. Reads that aligned within the deleted regions in the irradiated line coincide with TAREs that are shared between many subtelomeric domains.

GAACCCTAAACCCTGAACCCTAAACCTAAAACACTAACCCTGAACCCCAAACCCTGAACCCTAAACCCTAAAACCTG
AACCCTAAACCGGAACCCTGAACCCTAAACCCTAAACCCTGAACCATAAACCCTGAACCCTAAACCTAAACCCTGA
ACCCTGAACCCTAAACCCTGAACCCTAAACCCTGAACCCTAAACCTAAACCCTAAACCTAAACCCTAAACCCTGAAA
CCCTAAACCCTGAACCCTAAACCCTGAACCCTAAACCCTAAACCCTGAACCCTAAACCCTGAAACCCTGAACCCTAA
ACCCTAAACCCTGAAACCCTGAAACCCTAAACCCTAAACCTAAACCCTGAACCCTAAACCCTGAACCCTAAACCCTG
AACCCTAAACCCTAAACCCTAAACCCTGAACCCTAAACCCTAAACCCTAAACCTAAACCCTGAACCCTAAACCCTAA
ACCCTAAAACCTGAACCCTGAACCCTGAAACCCTGAACCCTAAACCCTAAACCCTGAACACTACACCCTAAACCCTG
AACCCTAAACCCTAAACCCTGAAACCCTGAACCCTAAACCCTAAACCCTGAACCCTAAACCCTAAACCCTAAACCCT
AAACCCTGAACCCTGAACCCTAAACCCTGAAACCTTGAAACAGTACTCCGTTCAACGTTCACTGCAGCATCAGATTT
TACTTTCACCCAAGTGCAAAAAGTCTCACGAGCGCCAGAGAAATCAAAAAAACATGACGATGGGTCATATTTTTGA
ATTATATCACTAGAAATCTTTAACGCACTAGTATAATCCGTTGTACTAAAATAGGATACCAATCGCTGATCACCTAC
AGTAAATATACCCATTTTTTGTAATGCTGCAATAACAAATTCCTTACCTGCCTTAGTACCTGCGGCTTCACCCGCTTC
AACACTAGCCTCAATACCCTTTTGAATCGCAGTTACTAAAGCGGCGCTTCTCCAAGCATAGATACCCAATCCTCCGA
ATAATCCAATACTTGCTGCAACACCTCCTAGCCCACACCCACACCTAAGACAACCTTTTTCTACTTTTTCTGCAAATG
ATTTGTCCTTCCTATCTTTTTCAATAATTTCTTGTATATTTTTGTCACGTTCTTCTTTACGTTTTTGGCGTTTATCTTTAA
TACGTTCTTCGTATTCTTCAAAACGTTGCGACGCTTGTCGATCGAAAGTTCCCTTCACTGATTTCATATCCGCATCAT
TATCATATTTAGGCAT

Figure 3.8. Assembled sequence showing the telomere healing event associated with the end of chromosome 1.This telomere healing event is shown schematically in Fig. 3.6A of the main text. The coding region of a *rif* gene is shown in black text, while the telomeric repeats are shown in blue.

ATGAAGGGATCTGGATCAGAAAAAAATGTATATCTTTCAAATAAAAATAAAGAAATTAATATGAACCAACAATCAG
ATAATAAAATGTGTGATGAATGTGATGATATGAATCAACCAGGAGATGTAAATAAAAATGACAAAACATCAAATG
ATCAAGCAAATTCAAGTGATTCTGATTGTGAGCCCTTACCATTTGGATTAAAACCTTCAGATTTAAATAGAAAAGTT
ACAGAAGAAGATTTAGAAAGAATGATAATAGAATTACCAGGAAAATTAGAAAGGAAAGATATGTATTTAATATGG
CATTATAGTCATTCTCTTTTGAGAGATAAATTTAATAAAATGAAAAGTTCGTTATGGAGTATTTGTGGGAAATTAGC
TCATGAACATAAGTTACCATTCAAAATTAAAATGAAGAAATGGTGGAAATGTTGTGGTCATGTTACAGATGAATTA
TTAATAAAAGAGCATGATGATTATAATTCTATATATAATTATATTAATAATGAATCATCAAGTCGTGAACAATTTCTT
ATATTTCTTAATATGATAAAGCATTCATGGACAACATTTACTATGGAGACTTTTATTAAATGTAAGATTTCTTTAGAA
AATAACATGAGAAATGTTAC**AGGGTTTAGGGTTCAGGGTTTAGGGTTTAGGGTTCAGGGTTTAGGGTTTAGGGTT**
TAGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGGTTCAGGGTTTAGGTTTAGGGTTTAGGGTTCAGGGT
TTAGGTTTAGGGTGTAGGGTTCAGGGTTTAGGGTTTAGGGTTCAGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGG
TTTAGGGTTCAGGGTTTAGGGTTCAGGGTTTAGGGTTCAGGGTTTAGGGTTCAGGGTTTAGGGTTCAGGGTTTAG
GGTTCAGGGTTAGGGTTCAGGGTTTAGGGTTCAGGGTTTAGGGTTCAGGGTTTTAGGGTTCAGGGTTCAGGGTTT
AGGGTTTAGGGTTCAGGGTTTAGGGTTCAGGGTTCAGGGTTTAGGGTTCAGGGTTCAGGGTTTAGGGTTTAGGGT
TCAGGGTTTAGGGTTCAGGGTTCAGGGTTTAGGGTTCAGGGTTTAGGTTCAGGGTTTAGGTTTTAGGTTTAGGGT
TCAGGGTTTAGGGTTTTAGGTTTAGGGTTCAGGGTTTAGGGTTCAGGGTTTAGGGTTTAGGGTTCAGGGTTCAGG

Figure 3.9. Assembled sequence showing the telomere healing event associated with the end of chromosome 2R. The telomere healing event is shown schematically in Fig. 3.6A of the main text. The coding region of Pf3D7_0221000 is shown in black text, while the telomeric repeats are shown in blue.

(Figure 3.10A). As predicted by this model, telomere healing is initiated at sites that display sequences complementary to the template region of telomerase RNA. In yeast, these sites almost always include GT, TG or CG di-nucleotides at the site where telomerase initiates synthesis of telomere repeat nine [25, 26]. In addition, the flanking region also plays a role in determining where the new telomere is synthesized, presumably by influencing how efficiently telomerase is recruited to the chromosome end [30]. In *S. cerevisiae*, stretches of TG repeats between 22 and 250 bp effectively recruit telomerase for healing, but stretches either longer or shorter are repaired much less efficiently [31-34]. Proximal enhancer sequences that can function as binding sites for proteins that associate with the telomerase complex can also increase the efficiency of telomere healing [30]. Sites of telomere healing in *P. falciparum* were previously identified within genes that are known to be commonly disrupted in cultured parasites, presumably because such deletion events provide a growth advantage in vitro [22, 24], and the ability of *P. falciparum* telomerase obtained from nuclear extracts to extend from specific sequences in vitro was also investigated [23]. The identification of 9 independent telomere healing events in our sequence data, including events that we induced with radiation exposure, allows us to investigate in more detail the properties at the site of a DSB that contribute to telomere healing in *P. falciparum* and whether they differ from sites that are instead repaired by HR.

The sequence of the telomerase RNA template region for several *Plasmodium* species was predicted bioinformatically by Chakrabarti et al. [35]. This template sequence is known to be species specific and, unlike the equivalent sequence in yeast, the *P. falciparum* sequence does not consist exclusively of As and Cs, and instead

includes two TG bases within the sequence 5'-ACCCTGAACCC-3'. Interestingly, given that the two major telomeric repeat sequences are 5'-TTCAGGG-3' and 5'-TTTAGGG-3', the G within the template sequence appears to specify either C or T. The sites where telomeric repeats were added to the chromosome ends for all nine telomere healing events identified in our sequence datasets are shown in Figure 3.10B. As can be easily discerned, the sites where repair was initiated in all nine sequences display the ability to anneal to the telomerase template region at precisely the same position, thus in all cases the first bases added to the new telomere were GGGTT. These data indicate telomere healing in *P. falciparum* likely involves the same mechanism described for yeast, yet the sequence of the telomerase RNA template explains why the site of telomere repeat addition diverges from the GT, TG or CG preference described for *S. cerevisiae*. Examination of the sequences upstream newly added telomere repeats did not identify any discernable motifs or compositional bias, indicating that either proximal enhancers like those identified in yeast do not exist in *P. falciparum*, or that recruitment of the telomerase complex is not influenced by primary sequence.

3.2.4 Choice of repair Pathway: HR vs telomere healing

In most extensively studied organisms, DSB repair in most regions of the genome results from competition between the two primary repair pathways, HR and NHEJ. Within subtelomeric regions of the genome, telomere healing can serve as a third potential repair pathway. The pathway that is ultimately chosen depends on

multiple factors and is often species specific [20, 21]. In yeast, the likelihood of telomere healing occurring at a specific break appears to depend on how efficiently the telomerase complex is recruited to the site of the break, which in turn depends on the DNA sequence immediately upstream of where the repair event occurs [20]. Telomerase can then initiate the addition of telomeric repeats when 3-6 bp of DNA anneals to the telomerase RNA template (Figure 3.10) [25]. Telomerase has been shown to be recruited by either CDC13 or the Ku70/80 complex [26, 36-38], and a proximal CDC13 binding sequence was shown to greatly increase the likelihood of a healing event at a particular chromosomal position [30]. Malaria parasites lack the above mentioned DNA repair proteins thus, what determines whether HR or telomere healing occurs at the site of a subtelomeric break is not known.

Previous work has shown that efficient HR within a non-subtelomeric region of the genome requires near complete sequence identity between the sequence surrounding the break and the template used for repair [15]. If this property also applies to HR within subtelomeric regions, the use of HR might be limited only to breaks that occur within stretches of sequence that are duplicated with near complete sequence identity elsewhere in the genome. To investigate this possibility, we identified the sequences immediately surrounding the breakpoints of the three products of HR that we obtained from our irradiated clone and performed BLAST searches to determine if these regions are duplicated elsewhere in the genome. Indeed in all three cases there were regions of perfect sequence identity surrounding the site of the recombination event. These stretches of 100% identity extend for 296, 127 and 27 bp within the

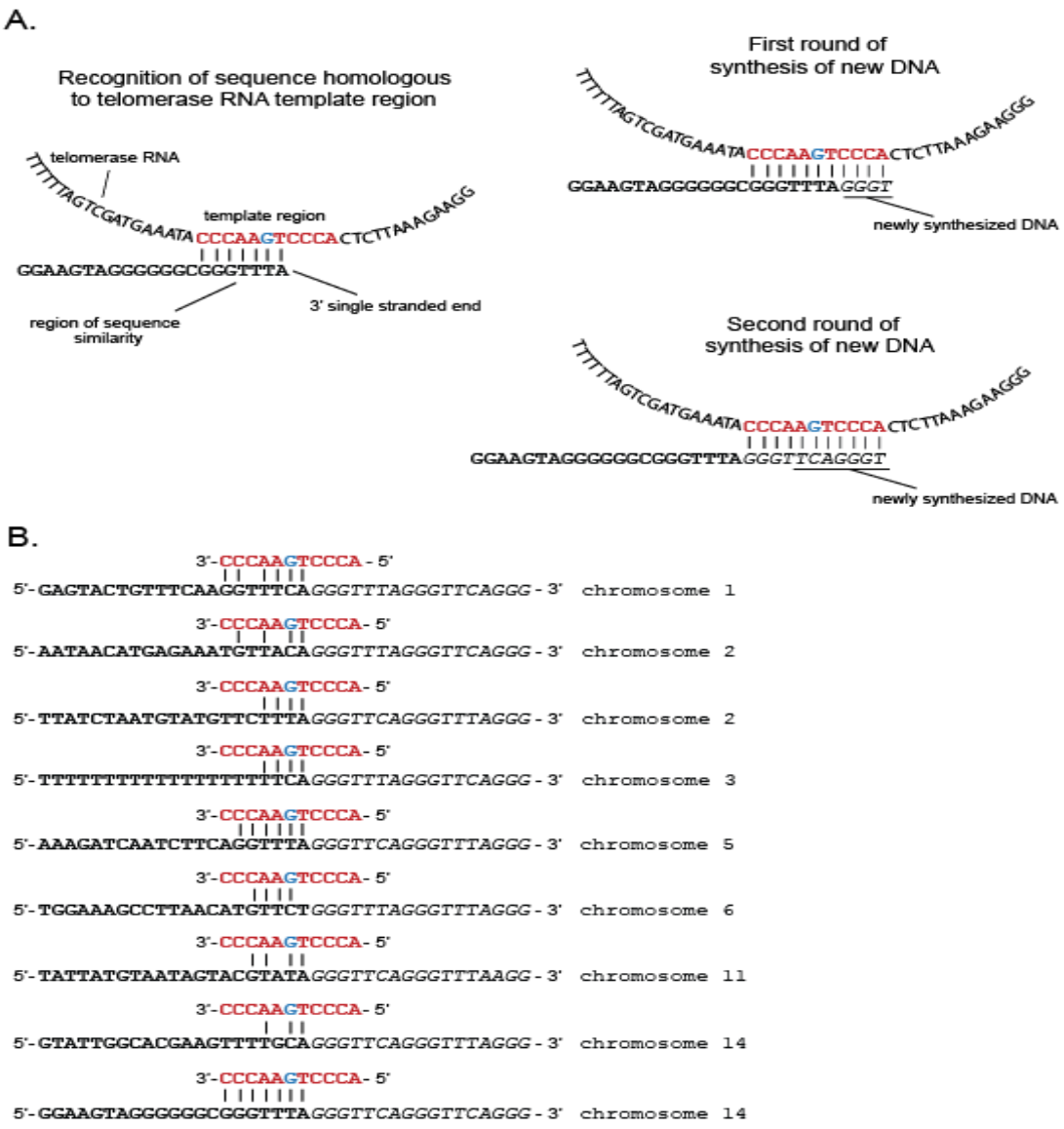


Figure 3.10. Sequence preference for telomere healing. A. Model for telomere repeat additions at the site of a DSB. After exonuclease resection, a 3' single strand overhang is revealed. Single stranded regions of sequence similarity can anneal to the template region of telomerase RNA (left). Telomerase activity can then extend by adding telomere repeat sequences directly to the chromosome end (right). Multiple rounds of repeat addition results in a lengthy telomere repeat region. B. The nine examples of telomere healing identified in this study. Bold sequence indicates the original region of the chromosome while italicized letters indicate the repeat sequences added by telomerase. The red sequence shows the template region of telomerase RNA and hypothetical annealing to the sequence where the healing event occurred. The blue "G" within the template indicates the bp that is thought to specify either C or T.[35]

sequences of the recombining chromosomes (Figure 3.11). In contrast, BLAST searches using stretches of sequence at the nine sites of telomere healing indicated that these regions are unique within the genome, with no identifiable stretches of sequence identity at the site of telomere addition. These data are consistent with a model in which HR is the dominant pathway of DSB repair throughout the genome, including within the subtelomeric regions. However, the strict requirement for extensive sequence identity prevents this pathway from repairing breaks that occur within unique sequences, thus allowing telomere healing to occur when sequences similar to the telomerase template RNA are exposed during resection of the DNA strand.

3.3 Discussion

Telomere healing appears to be a conserved mechanism for stabilizing chromosomes in the event of a DSB that occurs within a subtelomeric domain. The basic machinery involved in telomere healing is likely shared between malaria parasites and model eukaryotes. In yeast, DNA resection from the site of the break is mediated by either the Exo 1 or DNA2/Sgs1 pathways, and orthologues to both of these enzymes have been identified encoded within the *P. falciparum* genome (PF3D7_0725000 and PF3D7_1010200 respectively). In addition, telomerase reverse transcriptase (PfTERT) has also been identified and displays the expected motifs, though with many insertions of stretches of basic amino acid sequences, as is often seen in *Plasmodium* proteins [39]. However, our work revealed some unique aspects to

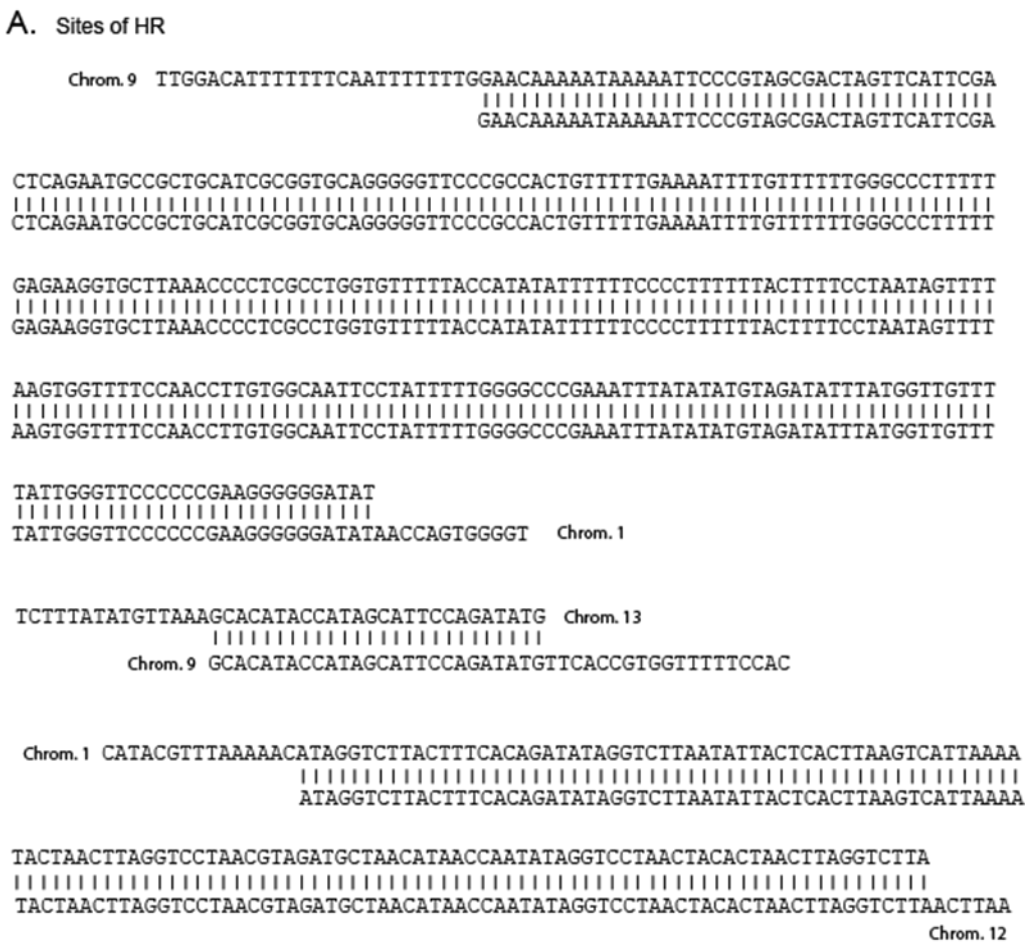


Figure 3.11. The sequences surrounding the breakpoints of the three examples of DSB repair by homologous recombination identified in this study. A and B show the recombination events that occurred within one of the subtelomeric regions of chromosome 12. The first (A) occurred within a region of 296 bp of sequence identity between chromosomes 9 and 1 while the second occurred within a region of 127 bp of sequence identity between chromosomes 1 and 12 (B). C shows a region of 27 of sequence identity spanning the breakpoint of the recombination event found within a subtelomeric region of chromosome 13, leading to the creation of a chimeric *var* gene.

telomere healing in *P. falciparum*. The strong preference for stretches of TG repeats at the site of telomere addition that is observed in yeast was not seen in *P. falciparum*, nor was any proximal enhancer sequence detected. The different sequence preference likely results from differences in the template region of telomerase RNA, which in yeast consists exclusively of AC bps, while in *P. falciparum* this sequence also includes a GT dinucleotide. The apparent lack of any preference in the sequence immediately flanking the site of repair might indicate more significant evolutionary divergence in how telomerase is recruited. Organism specific proteins at telomeres have been characterized in model organisms and African Trypanosomes and indicate there can be significant evolutionary divergence [40]. Proteins known to play important roles in telomerase function and telomere stabilization in other organisms, such as TRF, Cdc13, Rad 52 or POT1, could not be identified in the *P. falciparum* genome using standard bioinformatics approaches and thus the recruitment, retention and function of telomerase in *P. falciparum* is likely to have unique elements such as the recently characterized protein PfTRZ. This protein was found to be a functional homolog to the transcription factor TFIIIA, yet it is associated with parasite telomeres and had a role in telomere maintenance [41].

The subtelomeric domains of the chromosomes of *P. falciparum* are of significant research interest due to the large multi-copy gene families that reside within these regions. The unique structure of the chromatin found here has been shown to play a role in regulating clonally variant expression thereby facilitating the process of antigenic variation and immune system avoidance [42]. The position of these large, semi-redundant gene families within the subtelomeric domains of most or all of

the chromosomes means that these genomic regions share significant blocks of sequence identity. In addition, the clustering of these regions at the nuclear periphery makes them prime substrates for HR in the event of a DSB [11]. This unique genomic organization, combined with the absence of canonical NHEJ, provides a simple mechanism that drives the generation of chimeric genes and thus the vast diversity of these gene families that is observed in the field. Indeed a chimeric *var* gene was readily generated over the course of the experiments described here.

While repair of DSBs within subtelomeric domains by HR is the likely source of diversity within the multi-copy gene families, HR appears to not be efficient when a DSB occurs within a sequence that diverges more than ~2% identity from any possible template for repair [15]. Indeed, the three examples of HR identified in this study all displayed long stretches of complete sequence identity between the two chromosomal regions involved in the recombination events (Figure 3.11). This finding is confirmed by other studies that observed similar, albeit somewhat shorter regions of sequence identity at sites of recombination within *var* genes [43-45]. Given the extensive sequence diversity within *var*, *rifin*, *stevor* and *Pfmc-2TM* genes, the chance that a randomly occurring DSB will occur precisely at a position with sufficient sequence identity to another position in the genome to serve as a template for HR are low. In the absence of efficient NHEJ, such DSBs would generally be lethal. However, telomere healing provides an alternative pathway for stabilizing DSBs that occur within subtelomeric domains thus enabling parasites to survive DNA damage within these regions and maintain genome integrity. Chromosomes that have undergone telomere healing could later undergo HR when a subsequent DSB occurs within sequence that shares identity

with a region of a full-length subtelomeric domain. The resulting gene conversion event would reestablish typical chromosomal structure, including a full complement of clonally variant gene copies as well as TAREs (Figure 3.12). When such events occur within the coding regions of the variant antigen encoding genes, new genes are created, as demonstrated by the HR event that occurred within the subtelomeric region of chromosome 13 described here (Figure 3.6B) and as has been observed in other studies [43-45]. Of note these recombination events retain reading frame and general gene structure. Evidence for telomere loss and potential healing events in field isolates indicate that this mechanism of chromosome stabilization occurs in naturally circulating parasites [46-48]. This provides a model for mitotic diversification of these important gene families using both telomere healing to stabilize chromosome ends and HR when breaks occur within areas of sequence identity, thereby creating new chimeric genes and restoring complete subtelomeric regions. Telomere healing therefore provides a complementary method to HR for preserving the structure of chromosome ends. Together these two pathways of DSB repair function to maintain genome integrity and chromosome stability in the absence of robust NHEJ and also drive the generation of diversity within the clonally variant, multi-copy gene families of *P. falciparum.*

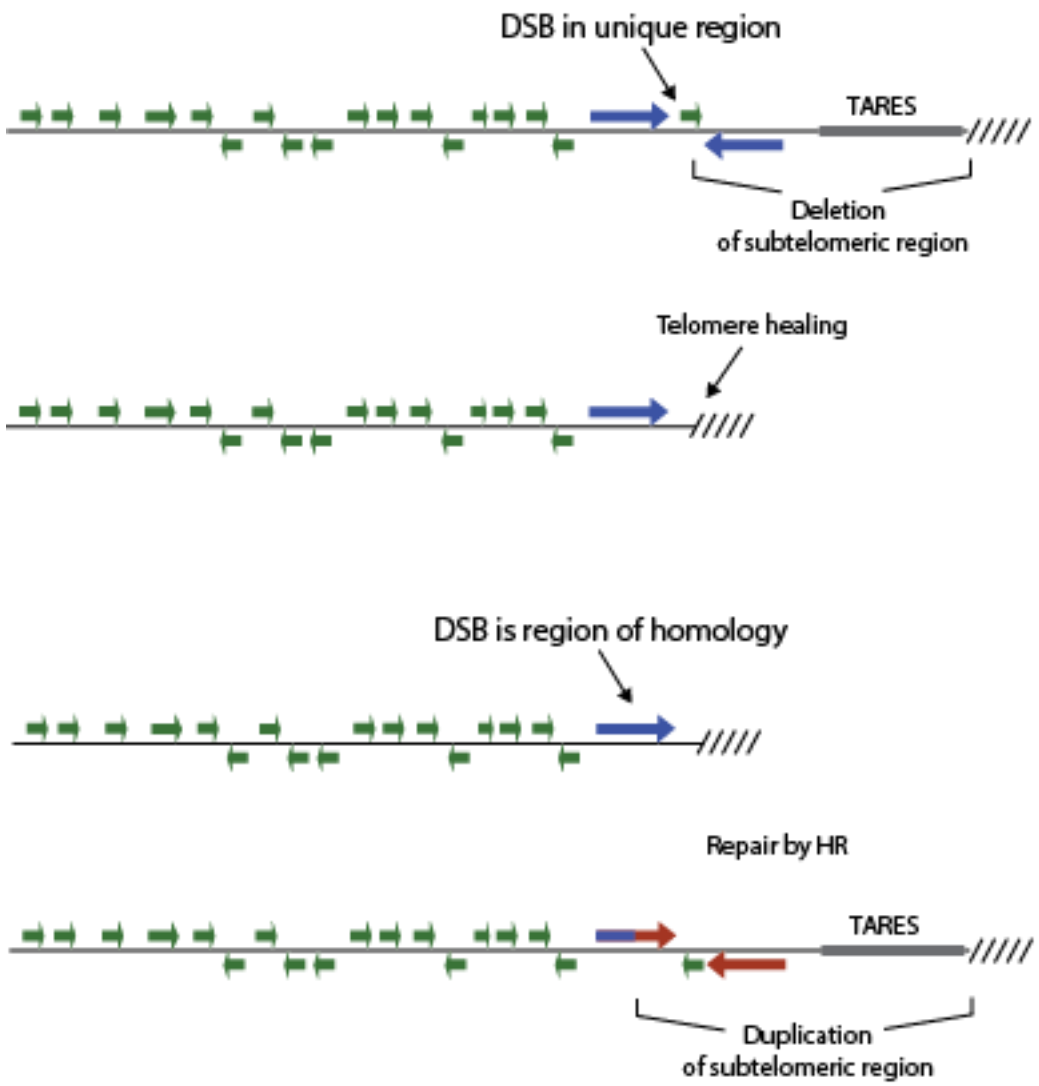


Figure 3.12. Model for the contribution of both telomere healing and homologous recombination in maintaining chromosome end stability in P. falciparum. A. The occurrence of a double strand break (DSB) at a site of unique sequence within a subtelomeric region is stabilized by telomere healing. This results in a substantial deletion of the subtelomeric domain, including members of multi-copy gene families and telomeres associated repeat elements (TARES). B. A subsequent DSB within a region that shares sequence identity with subtelomeric regions from other chromosomes can be repaired by homologous recombination (HR) leading to reestablishment of the normal subtelomeric structure, including a full complement of multi-copy genes and TARES. Repair by HR can also result in chimeric genes, thereby contributing to the generation of diversity within the multi-copy gene families.

3.4 Materials and Methods

Parasite culture: *P. falciparum* parasites were cultivated at 5% hematocrit in RPMI 1640 medium (Corning Life Sciences, Tewksbury, MA), 0.5% albumax II (Invitrogen), 0.25% sodium bicarbonate, and 0.1 mg/ml gentamicin. Cultures were maintained at 37 °C in an atmosphere of 5% oxygen, 5% carbon dioxide, and 90% nitrogen. Individual clonal lines were obtained by limiting dilution [49].

Parasite irradiation: Irradiation was administered using a Rad Source 2000 irradiator set at 160 kV/25 mA. To generate random DSBs, 3D7 parasites seeded at 0.5% were exposed to 100 Gy X-ray irradiation three times consecutively, being allowed to recover to normal growth between each irradiation exposure. The degree of subtelomeric damage was preliminarily assessed after each round of irradiation by assaying for *var* gene deletion by Q-PCR using gDNA as template and the *var* specific PCR primer set described previously by Salanti and colleagues [50]. The subclone chosen for whole genome sequencing displayed loss of three subtelomeric *var* clusters using this assay.

Genomic DNA Isolation: 100 mls of cultured parasites at 5-8% parasitemia were harvested for isolation of genomic DNA. DNA was isolated and purified using phenol-chloroform extraction followed by ethanol precipitation as previously described [51].

SMRTbell library preparation: The SMRTbell libraries were prepared as previously described [28] using SMRTbell Template Prep Kit 1.0 (Pacific Biosciences) and following the standard 20 kb template preparation using the BluePippin Size-Selection System protocol (Pacific Biosciences). Briefly, parasite DNA was sheared twice for 1

min at 5300 rpm in an Eppendorf 5424 centrifuge using a g-TUBE (Covaris) followed by damage repair, end repair and ligation of SMARTbell adapters. Unligated DNA was digested with exonucleases and the libraries were size-selected using a BluePippen pulsed-field gel electrophoresis instrument (Sage Science) to isolate fragments greater than 15 kb. Library concentration measured with the Qubit fluorometer dsDNA BR Assay kit (Life Technologies) and fragment length distributions were generated using 2200 TapeStation (Agilent). Sequencing primer and P6 polymerase were annealed to the libraries according to manufacturer's protocols (Pacific Biosciences) and done with P6-C5 chemistry and v3 SMRT Cells on an RSII instrument at Weill Cornell Medicine.

Genome sequencing and analysis: Pacific Biosciences RSII-based single molecule sequencing was used to prepare long read datasets suitable for accurate assembly of the subtelomeric regions of the parasite's genome. Sequencing libraries were produced, using the PacBio 20kb library preparation protocol, from high molecular weight gDNA obtained from clonal parasite lines. DNA concentration was measured by Qubit (Life Technologies) and fragment length distributions of starting material and final library were generated by Tapestation 2200 (Agilent). Six RSII SMRT Cells were used for the non-irradiated clone library while the irradiated clone library was sequenced with eight SMRTCells. Filtered sequence data were assembled using HGAP 2.0 with Quiver polishing. This Celera Assembler based de novo assembly approach produced genome sizes of 23.2 Mb and 23.3 Mb for the non-irradiated and irradiated clones, respectively. The largest contig was 3.2 Mb and assembled contig N50 values were of 1.5 Mb for both lines. The non-irradiated line generated 29

polished contigs, while the irradiated line produced 24 polished contigs. The average coverage for each assembly was 130X and 100X for the non-irradiated and irradiated lines, respectively, and the average consensus concordance for both assemblies was greater than 0.9975; therefore, only 2.5 bases in a thousand would be assigned incorrectly.

In order to identify recombination break points, telomere additions, or telomere deletions, each assembly was first aligned to a 3D7 genomic sequence from PlasmoDB using Mauve [28]. Using the Mauve alignment, chromosomes were assigned to each polished contig in both assemblies. 30 kb from the right and left end of each chromosome from the irradiated and non-irradiated polished assemblies were then aligned to the PlasmoDB reference genome using BLAST [52]. Recombination break points and telomere additions were determined via manual inspection of the tabulated BLAST output. Each break point was also confirmed using NCBI blastn online tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

To validate structural variants (SV) seen in comparisons between the PlasmoDB 3D7 reference and our assemblies, reads from the datasets used for assemblies of irradiated and non-irradiated clones were aligned to the Plasmo DB reference using BWA-MEM (version 0.7.13-r1126; arXiv:1303.3997) with default parameters. The resulting alignments were processed with SAMtools (version 2.6.32-279.el6.x86_64) [53] and visualized with IGV (version 2.3.92) for manual inspection of loci with predicted SVs. Analyses of the sequences upstream of the loci at which telomere repeats were added

were performed using MEME software suite and did not identify discernable motifs shared between any of the sequences.

References

1. Cousins, S. Targets to reduce malaria unlikely to be met, says WHO. *BMJ* **355**, i6668 (2016).
2. Miller, L.H., Baruch, D.I., Marsh, K. & Doumbo, O.K. The pathogenic basis of malaria. *Nature* **415**, 673-679 (2002).
3. Golenser, J. et al. Correlation between destruction of malarial parasites by polymorphonuclear leucocytes and oxidative stress. *Free Radic.Res.Commun.* **17**, 249-262 (1992).
4. Schnappinger, D. et al. Transcriptional Adaptation of Mycobacterium tuberculosis within Macrophages: Insights into the Phagosomal Environment. *J.Exp.Med.* **198**, 693-704 (2003).
5. Gopalakrishnan, A.M. & Kumar, N. Antimalarial action of artesunate involves DNA damage mediated by reactive oxygen species. *Antimicrob.Agents Chemother.* **59**, 317-325 (2015).
6. Gardner, M.J. et al. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419**, 498-511 (2002).
7. Figueiredo, L.M., Pirrit, L.A. & Scherf, A. Genomic organisation and chromatin structure of Plasmodium falciparum chromosome ends. *Mol.Biochem.Parasitol.* **106**, 169-174 (2000).
8. Chookajorn, T. et al. Epigenetic memory at malaria virulence genes. *Proc.Natl.Acad.Sci.U.S.A* **104**, 899-902 (2007).
9. Lopez-Rubio, J.J. et al. 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Mol.Microbiol.* **66**, 1296-1305 (2007).
10. Jiang, L. et al. PfSETvs methylation of histone H3K36 represses virulence genes in Plasmodium falciparum. *Nature* **499**, 223-227 (2013).
11. Freitas-Junior, L.H. et al. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. *Nature* **407**, 1018-1022 (2000).
12. Bhattacharyya, M.K., Norris, D.E. & Kumar, N. Molecular players of homologous recombination in protozoan parasites: implications for generating antigenic variation. *Infect.Genet.Evol.* **4**, 91-98 (2004).
13. Mladenov, E. & Iliakis, G. Induction and repair of DNA double strand breaks: the increasing spectrum of non-homologous end joining pathways. *Mutat.Res.* **711**, 61-72 (2011).
14. LaRocque, J.R. & Jasin, M. Mechanisms of recombination between diverged sequences in wild-type and BLM-deficient mouse and human cells. *Mol.Cell Biol.* **30**, 1887-1897 (2010).
15. Kirkman, L.A., Lawrence, E.A. & Deitsch, K.W. Malaria parasites utilize both homologous recombination and alternative end joining pathways to maintain genome integrity. *Nucleic Acids Res.* **42**, 370-379 (2014).
16. Singer, M. et al. Zinc finger nuclease-based double-strand breaks attenuate malaria parasites and reveal rare microhomology-mediated end joining. *Genome Biol.* **16**, 249 (2015).
17. Bhattacharyya, M.K. & Kumar, N. Identification and molecular characterisation of DNA damaging agent induced expression of Plasmodium falciparum recombination protein PfRad51. *Int.J.Parasitol.* **33**, 1385-1392 (2003).

18. Roy, N. et al. Dominant negative mutant of Plasmodium Rad51 causes reduced parasite burden in host by abrogating DNA double-strand break repair. *Mol.Microbiol.* **94**, 353-366 (2014).
19. Badugu, S.B. et al. Identification of Plasmodium falciparum DNA Repair Protein Mre11 with an Evolutionarily Conserved Nuclease Function. *PLoS.ONE.* **10**, e0125358 (2015).
20. Pennaneach, V., Putnam, C.D. & Kolodner, R.D. Chromosome healing by de novo telomere addition in Saccharomyces cerevisiae. *Mol.Microbiol.* **59**, 1357-1368 (2006).
21. Ribeyre, C. & Shore, D. Regulation of telomere addition at DNA double-strand breaks. *Chromosoma* **122**, 159-173 (2013).
22. Mattei, D. & Scherf, A. Subtelomeric chromosome instability in Plasmodium falciparum: short telomere-like sequence motifs found frequently at healed chromosome breakpoints. *Mutat.Res.* **324**, 115-120 (1994).
23. Bottius, E., Bakhsis, N. & Scherf, A. Plasmodium falciparum telomerase: de novo telomere addition to telomeric and nontelomeric sequences and role in chromosome healing. *Mol.Cell Biol.* **18**, 919-925 (1998).
24. Scherf, A. & Mattei, D. Cloning and characterization of chromosome breakpoints of Plasmodium falciparum: breakage and new telomere formation occurs frequently and randomly in subtelomeric genes. *Nucleic Acids Res.* **20**, 1491-1496 (1992).
25. Putnam, C.D., Pennaneach, V. & Kolodner, R.D. Chromosome healing through terminal deletions generated by de novo telomere additions in Saccharomyces cerevisiae. *Proc.Natl.Acad.Sci.U.S.A* **101**, 13262-13267 (2004).
26. Stellwagen, A.E., Haimberger, Z.W., Veatch, J.R. & Gottschling, D.E. Ku interacts with telomerase RNA to promote telomere addition at native and broken chromosome ends. *Genes Dev.* **17**, 2384-2395 (2003).
27. Eid, J. et al. Real-time DNA sequencing from single polymerase molecules. *Science* **323**, 133-138 (2009).
28. Vembar, S.S. et al. Complete telomere-to-telomere de novo assembly of the Plasmodium falciparum genome through long-read (>11 kb), single molecule, real-time sequencing. *DNA Res.* **23**, 339-351 (2016).
29. Bonetti, D., Martina, M., Clerici, M., Lucchini, G. & Longhese, M.P. Multiple pathways regulate 3' overhang generation at S. cerevisiae telomeres. *Mol.Cell* **35**, 70-81 (2009).
30. Obodo, U.C. et al. Endogenous Hot Spots of De Novo Telomere Addition in the Yeast Genome Contain Proximal Enhancers That Bind Cdc13. *Mol.Cell Biol.* **36**, 1750-1763 (2016).
31. Hirano, Y. & Sugimoto, K. Cdc13 telomere capping decreases Mec1 association but does not affect Tel1 association with DNA ends. *Mol.Biol.Cell* **18**, 2026-2036 (2007).
32. Hirano, Y., Fukunaga, K. & Sugimoto, K. Rif1 and rif2 inhibit localization of tel1 to DNA ends. *Mol.Cell* **33**, 312-322 (2009).
33. Zhang, W. & Durocher, D. De novo telomere formation is suppressed by the Mec1-dependent inhibition of Cdc13 accumulation at DNA breaks. *Genes Dev.* **24**, 502-515 (2010).
34. Negrini, S., Ribaud, V., Bianchi, A. & Shore, D. DNA breaks are masked by multiple Rap1 binding in yeast: implications for telomere capping and telomerase regulation. *Genes Dev.* **21**, 292-302 (2007).
35. Chakrabarti, K. et al. Structural RNAs of known and unknown function identified in malaria parasites by comparative genomics and RNA analysis. *RNA.* **13**, 1923-1939 (2007).

36. Pennock, E., Buckley, K. & Lundblad, V. Cdc13 delivers separate complexes to the telomere for end protection and replication. *Cell* **104**, 387-396 (2001).
37. Fisher, T.S., Taggart, A.K. & Zakian, V.A. Cell cycle-dependent regulation of yeast telomerase by Ku. *Nat.Struct.Mol.Biol.* **11**, 1198-1205 (2004).
38. DuBois, M.L., Haimberger, Z.W., McIntosh, M.W. & Gottschling, D.E. A quantitative assay for telomere protection in Saccharomyces cerevisiae. *Genetics* **161**, 995-1013 (2002).
39. Figueiredo, L. & Scherf, A. Plasmodium telomeres and telomerase: the usual actors in an unusual scenario. *Chromosome.Res.* **13**, 517-524 (2005).
40. Lira, C.B., Giardini, M.A., Neto, J.L., Conte, F.F. & Cano, M.I. Telomere biology of trypanosomatids: beginning to answer some questions. *Trends Parasitol.* **23**, 357-362 (2007).
41. Bertschi, N.L. et al. Malaria parasites possess a telomere repeat-binding protein that shares ancestry with transcription factor IIIA. *Nat.Microbiol.* **2**, 17033 (2017).
42. Cortes, A. & Deitsch, K.W. Malaria Epigenetics. *Cold Spring Harb.Perspect.Med.* (2017).
43. Sander, A.F. et al. DNA secondary structures are associated with recombination in major Plasmodium falciparum variable surface antigen gene families. *Nucleic Acids Res.* **42**, 2270-2281 (2014).
44. Claessens, A. et al. Generation of antigenic diversity in Plasmodium falciparum by structured rearrangement of Var genes during mitosis. *PLoS.Genet.* **10**, e1004812 (2014).
45. Bopp, S.E. et al. Mitotic evolution of Plasmodium falciparum shows a stable core genome but recombination in antigen families. *PLoS.Genet.* **9**, e1003293 (2013).
46. Dharia, N.V. et al. Genome scanning of Amazonian Plasmodium falciparum shows subtelomeric instability and clindamycin-resistant parasites. *Genome Res.* **20**, 1534-1544 (2010).
47. Murillo, S.C. et al. Deletion of Plasmodium falciparum Histidine-Rich Protein 2 (pfhrp2) and Histidine-Rich Protein 3 (pfhrp3) Genes in Colombian Parasites. *PLoS.ONE.* **10**, e0131576 (2015).
48. Biggs, B.A., Kemp, D.J. & Brown, G.V. Subtelomeric chromosome deletions in field isolates of Plasmodium falciparum and their relationship to loss of cytoadherence in vitro. *Proc.Natl.Acad.Sci.U.S.A* **86**, 2428-2432 (1989).
49. Kirkman, L.A., Su, X.Z. & Wellems, T.E. Plasmodium falciparum: isolation of large numbers of parasite clones from infected blood samples. *Exp.Parasitol.* **83**, 147-149 (1996).
50. Salanti, A. et al. Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering Plasmodium falciparum involved in pregnancy-associated malaria. *Molecular Microbiology* **49**, 179-191 (2003).
51. Swamy, L., Amulic, B. & Deitsch, K.W. Plasmodium falciparum var gene silencing is determined by cis DNA elements that form stable and heritable interactions. *Eukaryot.Cell* **10**, 530-539 (2011).
52. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J.Mol.Biol.* **215**, 403-410 (1990).
53. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* **25**, 2078-2079 (2009).

Chapter 4: Implications of DNA repair pathways on parasite fitness

Plasmodium falciparum possesses a unique genome that does not follow the canonical organization and mechanisms of other model eukaryotes. The parasite genome has a high A+T composition with enrichment of low complexity regions in protein-coding genes[1] and though haploid, relies on HR for DSB repair while displaying recombinant events not characteristically allowed by homologous repair[2, 3] indicating inherent plasticity of repair pathways. In fact, a degree of mutational variability is likely both evolutionarily and individually selected for in response to the host immune system and antimalarial drugs. However, this mutability cannot be explained by NHEJ, traditionally thought to be more mutagenic[4], since *Plasmodium* lacks the pathway's essential protein components[3, 5]. So far we have primarily investigated DSB repair in malaria parasites but single strand DNA repair pathways are also essential to maintain genome stability. In this chapter we will discuss the pathways of DNA repair: presenting evidence from model systems and divergences found in *Plasmodium* and integrating the data presented in Chapters 2 and 3. We also discuss how these findings illustrate how the parasite's unique genome is maintained by the unique aspects of its DNA damage pathways and evolutionarily skew the species toward mutagenic variability that enable antigenic variation and sustain infection in diverse hosts.

4.1 Single Strand DNA Repair and its Implications for Drug Targets and Drug Resistance

Base Excision Repair (BER) is likely the most used DNA repair pathway in cells and repairs alkylated, oxidized or deaminated bases.[6] DNA polymerases δ and β are used in BER to repair AP (apurinic/apyrimidinic) sites caused by spontaneous loss of bases or excision of damaged bases.[7] In humans and mammalian cells, after base excision the single nucleotide gap is filled in by DNA polymerase β with a single nucleotide patch.[7] This pol β dependent pathway that removes and repairs one nucleotide is called short-patch repair.[8] But experimental evidence has failed to detect this short-patch BER repair in *P. falciparum*.[7] Despite a *P. falciparum* DNA polymerase β-like enzyme being purified from parasite extract, only long-patch BER activity was detected.[9] Long-patch is the other BER pathway where an AP endonuclease cleaves the AP site and generates a repair patch of more than one nucleotide. Polymerase δ generates a repair patch of up to seven nucleotides, this activity in yeast and mammalian cells is dependent on proliferating cell nuclear antigen (PCNA) associated DNA.[10] *P. falciparum* Pol δ has been sequenced[11] and characterized.[12] While its activity is not PCNA dependent, the highest enzymatic processing of Pol δ was detected when PCNA was added[12]. These findings are striking considering that in both prokaryotes and eukaryotes the majority of BER is performed by the short patch mechanism.[13, 14] The parasite's mammalian host mainly uses short BER and this suggests a major difference between mammalian and *Plasmodium* DNA repair. Evidence suggests that *Plasmodium* long-patch repair is not PCNA dependent or parasite PCNA may be structurally different, while PCNA is essential to mammalian

long patch repair.[7] If the parasite PCNA is functionally different, this an attractive potential antimalarial target. Similarly, the *P. falciparum* DNA polymerase β-like enzyme exhibits different sensitivities to known inhibitors of mammalian DNA Pol β which would allow for the development of drugs to specifically target the *Plasmodium* enzyme.[9]

Nucleotide Excision Repair (NER) removes bulky and cross-linked DNA adducts that can distort the double helix structure.[15] NER repair was found to be defective in two drug-resistant parasite strains[16], suggesting the potential for an increase in mutation rate that could lead to the acquisition of drug-resistant mutations. Almost all the components of the NER machinery are found to be encoded in the *Plasmodium* genome, though the overall sequence identity between *P.falciparum,* yeast and human was low. Several components had large insertions of asparagine or aspartic acid residues[15]; these evolutionarily divergent features could have a species-specific functions linked to altered function in drug-resistant strains. Mismatch repair (MMR) was also found to have defective function in two ARMD parasites, W2 and DD2.[17] These findings give support to the idea that altered functionality of these repair pathways can lead to mutations conferring drug resistance.

4.2 *Plasmodium* Species-Specific Translesion Synthesis Machinery

The spontaneous formation of AP sites has been estimated at a rate of 10,000 sites per cell per day.[18] With depurination so frequent, the high A+T content of the *P. falciparum* genome[5] indicates that the parasites would be repairing AP sites frequently. If an AP site cannot be repaired by BER or another DNA repair pathway,

it can cause replication fork stalling. This stalling occurs because classical DNA polymerases such as Pol δ and Pol ε are unable to efficiently incorporate deoxynucleotides opposite DNA damage. Without a means to overcome these replication blocks, the replication fork collapses causing a dangerous DSB which can lead to chromosome rearrangement and cell death. To inhibit this outcome, eukaryotes have mechanisms to tolerate completion of replication in genomes with unrepaired DNA lesions. Translesion replication, where a nascent strand extension is promoted by specialized DNA polymerases, called translesion synthesis polymerases, is one such mechanism.[19] In *S. cerevisiae* the translesion synthesis polymerases are Pol ζ, Rev1 and Pol η, encoded by genes *REV3, REV1* and *RAD30*, respectively.[20] DNA Pol ζ is required for mutagenic bypass of DNA lesions in eukaryotes, but as it is a lower fidelity polymerase it incorporates endogenous and DNA-damage induced point mutations during translesion synthesis.[21] Mammalian REV3L and REV7 proteins constitute the two subunits of mammalian Pol ζ, with REV3L being the catalytic subunit.[21] REV3L has a polymerase domain and two REV7 binding domains, REV7 is necessary for Pol ζ activity *in vitro* and is needed as a bridge protein for interaction with REV1.[22] Mammalian REV1 is a scaffold protein that has a domain that binds with REV7 to allow interaction with Pol ζ. REV1 also interacts with DNA polymerase (Pol η, κ, or ι) which may facilitate multi-polymerase TLS.[23] Deletion of *REV3L* leads to death during embryogenesis for mice.[24] Mammalian cells in culture also cannot survive *REV3L* deletion as chromosomal DNA breaks accumulate.[25] Mammalian cells lacking REV1 are hypersensitive to UV irradiation but the DNA damage tolerance seems not to require the polymerase catalytic domain

of REV1. Rather the damage tolerance properties are granted by the protein-protein interactions between REV1 and REV7 in Pol ζ and other Y family DNA polymerases.[26]

In contrast, *S. cerevisiae* Pol ζ (REV3 and REV7) is not necessary for viability and genomic DNA replication in cells. *REV3* deletion however does produce a moderate increase in sensitivity to UV irradiation and some chemical DNA-damaging agents such as MMS.[24] This may be correlated with the fact that REV3 lacks 3' to 5' exonuclease activity and has relatively low-fidelity.[27] The most dramatic consequence of REV3 deletion is the significant decrease (by 90% or more) of base pair substitutions and frame-shift mutations caused by UV, gamma irradiation and MMS in yeast[28] and reduces the frequency of frameshift mutations by about half [29]. This would indicate that in yeast Pol ζ facilitates most translesion synthesis or mutagenic bypass events which would otherwise lead to stalled or collapsed replication forks. This provides a survival advantage but the cost of translesion DNA synthesis is the likelihood of mutagenesis during bypass of DNA damage. REV1 has an essential role in mutagenesis in yeast as it allows for association with ubiquitinated PCNA through its ubiquitin binding motifs.[30] Ubiquitinated PCNA (Ub-PCNA) is necessary for damage-induced mutagenesis as it causes the disassociation of DNA replication polymerases and the association of damage bypass polymerases.[31] REV1 binding with REV3 and REV7 allows for association with Ub-PCNA and enables Pol ζ to insert bases opposite DNA damage.[30] While TLS polymerases ensure the genome is maintained, the mutations introduced could contribute to a drug resistant phenotype

over time[16]. These polymerases are so low-fidelity, that about 50% of TLS events in mammalian cells are potentially mutagenic.[32]

Translesion synthesis can occur using other DNA polymerases, Pol η in yeast as well as Pol κ and ι in mammalian cells, while *Plasmodium* species lacks these other translesion synthesis polymerases[1] (orthologs searched for using PlasmoDB). This would indicate dependence on Pol ζ for bypass of DNA lesions during replication cycles. When examining orthology across *Plasmodium* species in PlasmoDB we see that primate and avian malaria species have Pol ζ , the catalytic subunit REV3, and REV1 while rodent malaria species lack these genes. It is unknown why the rodent malaria parasites *P. berghei, P. yoelii* and *P. chaubaudi* lack these proteins, therefore indicating they would be unable to perform translesion synthesis when needed during DNA replication. The evolutionary history of *Plasmodium* species has been heavily debated, a previous theory that *P. falciparum* was monophyletic with the avian subgroup and share a recent avian progenitor[33] has been disproven. More recent analyses of the three subgroups: avian, rodent and primate, support a mammalian clade and *P. falciparum* is no longer grouped with avian parasites.[34-36] There is a high level of conservation of genes between *P. falciparum* and rodent parasites in what are called syntenic regions of the genome.[37] The lack of *REV3* and *REV1* genes in the rodent genomes is called an intrasyntenic indel which is a disruption of synteny blocks[38]. The deletion of these translesion synthesis polymerases at the expected locus in *Plasmodium* rodent species could be caused by an early deletion event during the separation of rodent from non-rodent *Plasmodium* branches.[34, 37]

To better understand this mechanistic loss we are currently creating REV3 and REV1 knockout *P. falciparum* transgenic lines using the CRISPR-Cas9 system. The pUF1-Cas9 plasmid and the pL6 plasmid containing a sgRNA targeting the *REV1* gene (PF3D7_0910500) or *REV3* gene (PF3D7_1037000) and 5' and 3' homology blocks of ~500bp to mediate deletion of the respective gene loci were modified from Ghorbal et al.[39] Both the pUF1-Cas9 and pL6-Rev3KO or pL6-Rev1KO plasmids were transfected into WT 3D7 parasites by using "DNA loaded" erythrocytes as previously described.[40] For stable transfections parasites were cultured with 40 ng/ml WR99210 (Jacobus Pharmaceuticals, Plainsboro, NJ, USA) and 1.5 μM DSM1 and cultivated at 5% hematocrit in RPMI 1640 medium (Corning Life Sciences, Tewksbury, MA), 0.5% albumax II (Invitrogen), 0.25% sodium bicarbonate, and 0.1 mg/ml gentamicin. Cultures were maintained at 37 °C in an atmosphere of 5% oxygen, 5% carbon dioxide, and 90% nitrogen. The Rev3KO transgenic parasite line has been successfully created and Rev1KO transgenic line is in progress. Preliminarily, the Rev3KO line is viable and does not appear to have an obvious growth defect though formal growth rate assays have not yet been performed. We hypothesize that both the Rev3 KO and Rev1 KO parasites will have increased sensitivity to DNA damage caused by X-ray irradiation as compared to wild type 3D7 parasites indicated by our irradiation sensitivity assay a described in Chapter 2. We also plan to examine these transgenic lines' ability to survive ssDNA damage caused by mitomycin C/cisplatin that causes intrastrand crosslinks and bleomycin which causes ssDNA breaks. This will allow us to examine how Pol ζ and REV1 functions influence tolerance of parasite DNA damage. Depending on these results we will

obtain whole genome sequences of the Rev3KO and Rev1KO parasites after inducing broad spectrum or site-specific DNA damage to determine how this affects the parasite's ability for mutagenesis and alters mechanisms of repair at specific sites. Genetic deletion of these genes has not yet been attempted in *P. falciparum* and this work will provide insight into translesion synthesis in malaria parasites and how these enzymes mediate parasite fitness within a vertebrate host.

4.3 Unique Aspects of Homology Mediated Repair in *P. falciparum*

The repair of DNA damage within the *P. falciparum* genome is dictated by two factors: the sequence surrounding the site of the DNA break or lesion and the stage at which the DSB occurs in the parasite lifecycle. These two factors determine which DNA repair pathway is most likely to be utilized at the site of the lesion/break. The *P. falciparum* species-specific proteins involved in the chosen pathway mediate if the repair site will preserve the original sequence or introduce mutagenesis, driving variation in the genome. Either option allows the parasite to maintain genome stability but the unusual repair systems of malaria parasites seems to select for a level of mutagenic variability that is beneficial to parasite fitness as mediated by host immunity[41] and antimalarials[42].

HR is the dominant repair pathway for DSBs in *P .falciparum* as evidenced by many studies examining DNA repair in this haploid organism.[3, 43, 44] In terms of other DSB repair methods, A-NHEJ is a rarely used mechanism as previously discussed in Chapter 1 and single-strand annealing (SSA) has not been documented in *P. falciparum.* SSA also requires Rad52[45, 46], which is a homology repair protein that

Plasmodium lacks[47]. Previous analysis of Rad51 function in *P. berghei* parasites by Roy et al. indicated that in PfRad51^{K134R} mutants, there was no functional HR pathway to repair DNA damage. This suggests that in the absence of homology-mediated repair there is no other DSB repair mechanism that can effectively compensate in *Plasmodium.*[43] However this theory is complicated by our research indicating that in *P. falciparum,* a Rad51 dominant negative phenotype does not correlate to a complete lack of parasite survival in response to irradiation. Our results in Chapter 2 suggest that *Plasmodium* parasites, while having dominant HR-mediated repair as evidenced by the hypersensitivity of Rad51DN parasites to X-ray irradiation, can effectively repair DSB by other repair mechanisms if given sufficient recovery time. It's possible that some parasites were unaffected by irradiation, though this population would be very small as we see complete clearance in the culture before regrowth. Our previous irradiation work examining irradiated clones indicated that all parasites that outgrew had encountered DNA damage when their genomes were analyzed for rearrangements and deletions.[48] To definitively determine if any parasites from the regrowth population were undamaged we could sequence clones examining either the core genome or subtelomeric regions for damage repair. We found the effect of the Rad51 loss of function mutation to be parasite stage specific and irradiation dose dependent. R51DN trophozoites exhibited the highest irradiation sensitivity as compared to WT trophozoites, with a 1.5 fold increase in the time to recovery. The greater resistance to irradiation of the trophozoite stage may allow for a clearer dominant negative phenotype as their ability to perform HR significantly affects their time to recovery. In contrast, since rings are haploid and thus cannot efficiently utilize HR, the

Rad51DN is expected to have a smaller relative effect. Therefore, the Rad51DN trophozoites should resemble the irradiation sensitivity of WT rings as the PfRad51^{K134R} mutation has inhibited the trophozoites from being able to perform HR repair. This prediction is consistent with our data (Figure 2.3). The parasite regrowth of R51DN irradiated parasites suggests that the DNA damage sustained by the parasites can be repaired by a yet unexplained mechanism of DNA repair or uncharacterized redundancies in the HR pathway. Designing a knockout of PfRad51 would best allow for analyzing redundant function in the HR repair pathway.

Examining HR repair in *P. falciparum* using previous research presents a model that facilitates homologous recombination that is Rad51-dependent. A DSB is sensed by PfMre11 and PfRad50, enzymes known to interact *in vitro.*[49] However, *Plasmodium* lacks the Xrs2 element of the complex, this important protein may have a functional ortholog but with an amino acid sequence that is not conserved. The 5' to 3' resection can occur with the Exo1, Sgs1, Top3, and Dna2 proteins which have all been identified bioinformatically in *Plasmodium.*[47] The 3' ssDNA ends can then be bound by a ssDNA binding protein called RPA in mammals that has three subunits. One subunit of RPA1 has been identified with two open reading frames in *P. falciparum*, a short truncated protein RPA1S and a longer protein, RPA1L.[50] PfRPA1S was found to not have any single strand exchange function and actively down regulates the single strand exchange (SSE) function of PfRPA1L, which can initiate SSE in the presence of PfRad51. Intriguingly, both PfRPA1L and PfRPA1S are differentially expressed throughout the parasite life cycle. PfRPA1L, which has functional SSE activity is always more highly expressed than PfRPA1S and is

predominantly expressed throughout the trophozoite and schizont stages and less so during the ring stage.[51] This makes sense since during ring and schizont stages HR repair cannot occur, while during the trophozoite stage there is the possibility for HR repair using a homologous template. Rad54 is a double-stranded DNA-dependent ATPase, another key player in HR and is present in multiple stages of the process.[52] While *Plasmodium* lacks Rad52[47], thought to be a necessary recruiter of Rad51[53], PfRad54 was found to accelerate homologous DNA single strand exchange in the presence of $CaCl_2$ and interacts with PfRad51.[50] This corroborates data that mammalian Rad51 requires Ca^{2+} as a cofactor to promote DNA strand exchange.[54] While *P. falciparum* lacks some of the canonical HR proteins found in yeast and mammals, the interaction and activation of *P. falciparum* HR homologues elucidates how HR- mediated repair can occur. Evidence has been found of Rad51-independent single strand exchange, the critical step that Rad51 is usually required for [46] and considering *rad51Δ* yeast cells are sensitive to DNA damage but viable[55, 56], this could be present in *P. falciparum.* The human Rad52 protein was found to catalyze D-loop formation in superhelical DNA and could promote homologous pairing and strand exchange using oligonucleotides. A precise stoichiometric complex of HsRad52 with ssDNA is needed to saturate the ssDNA for optimal strand exchange.[57] As well, a *Rad51* deletion yeast mutant was found to stimulate gene conversion between homologous chromosomes but required Rad52.[58] However no PfRad52 ortholog has been identified[47] though a low homology ortholog of BRCA2 was discovered in *P. falciparum*[47] which has been found to regulate Rad51 and assist nucleoprotein filament formation in mammalian cells.[59] Though BRCA2-dependent homologous

recombination has been observed in mammalian cells, this cannot occur without Rad51 strand exchange.[60] Lastly, when the N-terminal domain of PfRad54 was expressed, it could not initiate SSE on its own but if the full-length PfRad54 were expressed, including the DNA-binding domain, it is possible that PfRad54 could have SSE activity.[50] This research presents possible Rad51-independent HR repair mechanisms through PfRad54 or a protein that has Rad52 function. Further characterization of the HR pathway and *Plasmodium* orthologs is needed to better understand possible pathway redundancies. Repair without a homologous template is still not understood in *Plasmodium* and ring survival of DNA damage may indicate a yet uncharacterized method of repair or identified proteins with unknown additional functions.

4.4 Specialized Repair in Subtelomeric Regions

For the subtelomeric regions of the genome we have shown in Chapter 3 that parasites use two competing DNA repair mechanisms to maintain genome integrity: homologous recombination and de novo telomere addition.[61] The subtelomeric domains of *P. falciparum* are unique and integral to understanding parasite biology, as the clonally variant gene families such as *var, rifin, stevor, Pfmc-2TM, FIKK* and *ACS* there are primary virulence determinants and encode proteins that are exposed on the surface of the infected erythrocyte. The most studied family is *var*, the main virulence factor of the disease. It encodes a polymorphic adhesion protein whose variant forms are called *Plasmodium falciparum* erythrocyte membrane protein 1 (*Pf*EMP1). The varied expression of these proteins enables antigenic variation, a mechanism which

allows the parasite to sustain long-term, chronic infections. These multicopy gene families within the subtelomeric domain are found in a specialized chromatin structure that is delineated by the histone modifications H3K9me3 and H3K36me3.[62, 63] It is theorized that HR is facilitated at these genes located in the subtelomeric region and internal clusters because of the unique histone modifications and chromatin structure at these sites.[41] We identified three recombination events in the irradiated *P. falciparum* genome, resulting from HR, with one creating a chimeric *var gene.* The three recombination events all occurred by the site of a DSB that had regions of perfect sequence identity surrounding it. Given the extensive sequence diversity at the subtelomeric regions, the chance that a randomly occurring DSB will occur precisely at a position with sufficient sequence identity to another position in the genome to serve as a template for HR are low. However, de novo telomere addition or telomere healing provides an alternative pathway for stabilizing DSBs that occur within subtelomeric domains thus enabling parasites to survive DNA damage within these regions and maintain genome integrity. This pathway has been described in model organisms and in *P. falciparum.*[64-66] Chromosomes that have undergone telomere healing could later undergo HR when a subsequent DSB occurs within a sequence that shares identity with a region of a full-length subtelomeric domain. The resulting gene conversion event would reestablish the typical chromosomal structure, including a full complement of clonally variant gene copies as well as TAREs preserving genomic stability. This method of telomere healing discussed here and in Chapter 3 is a third pathway of DNA repair besides HR and A-NHEJ that has been characterized and studied in *P. falciparum*.

4.5 The Mutagenic Variability of the *Plasmodium* genome

Genome plasticity is essential to the malaria parasite as a way to drive antigenic variation and promote DNA polymorphisms which contribute to drug resistance and immune evasion. The rapid emergence of anti-malarial drug resistant phenotypes is thought to be linked to the increased mutation ability of some parasite strains (mutator phenotype), called accelerated resistance to multiple drugs (ARMD).[42] Polymorphisms in DNA repair genes involved in BER, MMR and NER pathways can impair repair function and cause mutagenesis[67]. Translesion synthesis polymerases allow for the bypass of DNA lesions at the cost of incorporating mutations into the sequence.[20] However, these polymerases are often essential as they can move past stalled replication forks, preventing collapse. Pol ζ is the only translesion polymerase encoded in the *Plasmodium* genome and therefore may be the only way to resolve certain DNA lesions.[1] However, rodent parasites lack Pol ζ and have no mutagenic bypass mechanism that has been characterized.[34] In the future, we plan to analyze the *P.falciparum* Pol ζ through genetic manipulation and elucidate the necessity of DNA damage polymerases for mutagenesis and parasite fitness. All the methods of DNA repair discussed above allow for stabilization of the genome while skewing towards mutagenic variability which can lead to development of drug resistance and facilitate antigenic variation. These repair pathways maintain the essential core genome while creating DNA polymorphisms in drug-resistance related genes[42, 67] and facilitating rearrangements in the subtelomeric regions[41, 61].

The work discussed here builds on knowledge of DNA damage repair pathways, the interaction and characterization of the proteins involved and how the conservation of these unique repair systems shapes the evolutionary lineage and fitness of the *Plasmodium* species to maintain a foothold in the vertebrate hosts of the world. It is necessary that we further our understanding of antigenic variation and the recombination events that drive it, how the haploid genome manages to repair itself within the confines of its eukaryotic machinery, and the evolutionary divergence of its pathways and how this influences pathogenesis.

References

1. Aravind, L., Iyer, L.M., Wellems, T.E. & Miller, L.H. Plasmodium biology: genomic gleanings. *Cell* **115**, 771-85 (2003).
2. Larocque, J.R. & Jasin, M. Mechanisms of recombination between diverged sequences in wild-type and BLM-deficient mouse and human cells. *Mol Cell Biol* **30**, 1887-97 (2010).
3. Kirkman, L.A., Lawrence, E.A. & Deitsch, K.W. Malaria parasites utilize both homologous recombination and alternative end joining pathways to maintain genome integrity. *Nucleic Acids Res* **42**, 370-9 (2014).
4. Betermier, M., Bertrand, P. & Lopez, B.S. Is non-homologous end-joining really an inherently error-prone process? *PLoS Genet* **10**, e1004086 (2014).
5. Gardner, M.J. et al. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419**, 498-511 (2002).
6. Morita, R. et al. Molecular mechanisms of the whole DNA repair system: a comparison of bacterial and eukaryotic systems. *J Nucleic Acids* **2010**, 179594 (2010).
7. Haltiwanger, B.M. et al. DNA base excision repair in human malaria parasites is predominantly by a long-patch pathway. *Biochemistry* **39**, 763-72 (2000).
8. Hakem, R. DNA-damage repair; the good, the bad, and the ugly. *EMBO J* **27**, 589-605 (2008).
9. Nunthawarasilp, P., Petmitr, S. & Chavalitshewinkoon-Petmitr, P. Partial purification and characterization of DNA polymerase beta-like enzyme from Plasmodium falciparum. *Mol Biochem Parasitol* **154**, 141-7 (2007).
10. Frosina, G. et al. Two pathways for base excision repair in mammalian cells. *J Biol Chem* **271**, 9573-8 (1996).
11. Ridley, R.G. et al. DNA polymerase delta: gene sequences from Plasmodium falciparum indicate that this enzyme is more highly conserved than DNA polymerase alpha. *Nucleic Acids Res* **19**, 6731-6 (1991).
12. Vasuvat, J. et al. Biochemical and functional characterization of Plasmodium falciparum DNA polymerase delta. *Malar J* **15**, 116 (2016).
13. Dianov, G., Bischoff, C., Piotrowski, J. & Bohr, V.A. Repair pathways for processing of 8-oxoguanine in DNA by mammalian cell extracts. *J Biol Chem* **273**, 33811-6 (1998).
14. Dianov, G., Price, A. & Lindahl, T. Generation of single-nucleotide repair patches following excision of uracil residues from DNA. *Mol Cell Biol* **12**, 1605-12 (1992).
15. Tajedin, L., Anwar, M., Gupta, D. & Tuteja, R. Comparative insight into nucleotide excision repair components of Plasmodium falciparum. *DNA Repair (Amst)* **28**, 60-72 (2015).
16. Trotta, R.F., Brown, M.L., Terrell, J.C. & Geyer, J.A. Defective DNA repair as a potential mechanism for the rapid development of drug resistance in Plasmodium falciparum. *Biochemistry* **43**, 4885-91 (2004).
17. Castellini, M.A. et al. Malaria drug resistance is associated with defective DNA mismatch repair. *Mol Biochem Parasitol* **177**, 143-7 (2011).
18. Lindahl, T. & Nyberg, B. Rate of depurination of native deoxyribonucleic acid. *Biochemistry* **11**, 3610-8 (1972).
19. Lange, S.S., Tomida, J., Boulware, K.S., Bhetawal, S. & Wood, R.D. The Polymerase Activity of Mammalian DNA Pol zeta Is Specifically Required for Cell and Embryonic Viability. *PLoS Genet* **12**, e1005759 (2016).

20. Waters, L.S. et al. Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. *Microbiol Mol Biol Rev* **73**, 134-54 (2009).
21. Shen, X. et al. REV3 and REV1 play major roles in recombination-independent repair of DNA interstrand cross-links mediated by monoubiquitinated proliferating cell nuclear antigen (PCNA). *J Biol Chem* **281**, 13869-72 (2006).
22. Kikuchi, S., Hara, K., Shimizu, T., Sato, M. & Hashimoto, H. Structural basis of recruitment of DNA polymerase zeta by interaction between REV1 and REV7 proteins. *J Biol Chem* **287**, 33847-52 (2012).
23. Murakumo, Y. et al. Interactions in the error-prone postreplication repair proteins hREV1, hREV3, and hREV7. *J Biol Chem* **276**, 35644-51 (2001).
24. Gan, G.N., Wittschieben, J.P., Wittschieben, B.O. & Wood, R.D. DNA polymerase zeta (pol zeta) in higher eukaryotes. *Cell Res* **18**, 174-83 (2008).
25. Lange, S.S. et al. Dual role for mammalian DNA polymerase zeta in maintaining genome stability and proliferative responses. *Proc Natl Acad Sci U S A* **110**, E687-96 (2013).
26. Ross, A.L., Simpson, L.J. & Sale, J.E. Vertebrate DNA damage tolerance requires the C-terminus but not BRCT or transferase domains of REV1. *Nucleic Acids Res* **33**, 1280-9 (2005).
27. McCulloch, S.D. & Kunkel, T.A. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res* **18**, 148-61 (2008).
28. Lawrence, C.W. & Maher, V.M. Mutagenesis in eukaryotes dependent on DNA polymerase zeta and Rev1p. *Philos Trans R Soc Lond B Biol Sci* **356**, 41-6 (2001).
29. Harfe, B.D. & Jinks-Robertson, S. DNA polymerase zeta introduces multiple mutations when bypassing spontaneous DNA damage in Saccharomyces cerevisiae. *Mol Cell* **6**, 1491-9 (2000).
30. Wood, A., Garg, P. & Burgers, P.M. A ubiquitin-binding motif in the translesion DNA polymerase Rev1 mediates its essential functional interaction with ubiquitinated proliferating cell nuclear antigen in response to DNA damage. *J Biol Chem* **282**, 20256-63 (2007).
31. Stelter, P. & Ulrich, H.D. Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. *Nature* **425**, 188-91 (2003).
32. Adar, S., Izhar, L., Hendel, A., Geacintov, N. & Livneh, Z. Repair of gaps opposite lesions by homologous recombination in mammalian cells. *Nucleic Acids Res* **37**, 5737-48 (2009).
33. Waters, A.P., Higgins, D.G. & McCutchan, T.F. Plasmodium falciparum appears to have arisen as a result of lateral transfer between avian and human hosts. *Proc Natl Acad Sci U S A* **88**, 3140-4 (1991).
34. Martinsen, E.S., Perkins, S.L. & Schall, J.J. A three-genome phylogeny of malaria parasites (Plasmodium and closely related genera): evolution of life-history traits and host switches. *Mol Phylogenet Evol* **47**, 261-73 (2008).
35. Davalos, L.M. & Perkins, S.L. Saturation and base composition bias explain phylogenomic conflict in Plasmodium. *Genomics* **91**, 433-42 (2008).
36. Perkins, S.L., Sarkar, I.N. & Carter, R. The phylogeny of rodent malaria parasites: simultaneous analysis across three genomes. *Infect Genet Evol* **7**, 74-83 (2007).
37. Kooij, T.W. et al. A Plasmodium whole-genome synteny map: indels and synteny breakpoints as foci for species-specific genes. *PLoS Pathog* **1**, e44 (2005).
38. Kooij, T.W., Janse, C.J. & Waters, A.P. Plasmodium post-genomics: better the bug you know? *Nat Rev Microbiol* **4**, 344-57 (2006).
39. Ghorbal, M. et al. Genome editing in the human malaria parasite Plasmodium falciparum using the CRISPR-Cas9 system. *Nat Biotechnol* **32**, 819-21 (2014).

40. Deitsch, K., Driskill, C. & Wellems, T. Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes. *Nucleic Acids Res* **29**, 850-3 (2001).
41. Kirkman, L.A. & Deitsch, K.W. Recombination and Diversification of the Variant Antigen Encoding Genes in the Malaria Parasite Plasmodium falciparum. *Microbiol Spectr* **2** (2014).
42. Rathod, P.K., McErlean, T. & Lee, P.C. Variations in frequencies of drug resistance in Plasmodium falciparum. *Proc Natl Acad Sci U S A* **94**, 9389-93 (1997).
43. Roy, N. et al. Dominant negative mutant of Plasmodium Rad51 causes reduced parasite burden in host by abrogating DNA double-strand break repair. *Mol Microbiol* **94**, 353-66 (2014).
44. Straimer, J. et al. Site-specific genome editing in Plasmodium falciparum using engineered zinc-finger nucleases. *Nat Methods* **9**, 993-8 (2012).
45. Ivanov, E.L., Sugawara, N., Fishman-Lobell, J. & Haber, J.E. Genetic requirements for the single-strand annealing pathway of double-strand break repair in Saccharomyces cerevisiae. *Genetics* **142**, 693-704 (1996).
46. Symington, L.S. Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol Mol Biol Rev* **66**, 630-70, table of contents (2002).
47. Lee, A.H., Symington, L.S. & Fidock, D.A. DNA repair mechanisms and their biological roles in the malaria parasite Plasmodium falciparum. *Microbiol Mol Biol Rev* **78**, 469-86 (2014).
48. Fortin, G.S. & Symington, L.S. Mutations in yeast Rad51 that partially bypass the requirement for Rad55 and Rad57 in DNA repair by increasing the stability of Rad51-DNA complexes. *EMBO J* **21**, 3160-70 (2002).
49. Badugu, S.B. et al. Identification of Plasmodium falciparum DNA Repair Protein Mre11 with an Evolutionarily Conserved Nuclease Function. *PLoS One* **10**, e0125358 (2015).
50. Gopalakrishnan, A.M. & Kumar, N. Opposing roles for two molecular forms of replication protein A in Rad51-Rad54-mediated DNA recombination in Plasmodium falciparum. *MBio* **4**, e00252-13 (2013).
51. Rider, S.D., Jr. et al. The protozoan parasite Cryptosporidium parvum possesses two functionally and evolutionarily divergent replication protein A large subunits. *J Biol Chem* **280**, 31460-9 (2005).
52. Tan, T.L., Kanaar, R. & Wyman, C. Rad54, a Jack of all trades in homologous recombination. *DNA Repair (Amst)* **2**, 787-94 (2003).
53. Sugawara, N., Wang, X. & Haber, J.E. In vivo roles of Rad52, Rad54, and Rad55 proteins in Rad51-mediated recombination. *Mol Cell* **12**, 209-19 (2003).
54. Mazina, O.M. & Mazin, A.V. Human Rad54 protein stimulates DNA strand exchange activity of hRad51 protein in the presence of Ca2+. *J Biol Chem* **279**, 52042-51 (2004).
55. Vishnevetskaia, O., Luchnik, A.N., Arutiunova, L.S. & Shestakov, S.V. [Disorder of the repair of DNA double-stranded breaks in radiosensitive mutants of Saccharomyces cerevisiae yeasts]. *Genetika* **19**, 26-32 (1983).
56. Moore, C.W. Responses of radiation-sensitive mutants of Saccharomyces cerevisiae to lethal effects of bleomycin. *Mutat Res* **51**, 165-80 (1978).
57. Kumar, J.K. & Gupta, R.C. Strand exchange activity of human recombination protein Rad52. *Proc Natl Acad Sci U S A* **101**, 9562-7 (2004).

58. Pohl, T.J. & Nickoloff, J.A. Rad51-independent interchromosomal double-strand break repair by gene conversion requires Rad52 but not Rad55, Rad57, or Dmc1. *Mol Cell Biol* **28**, 897-906 (2008).
59. Davies, A.A. et al. Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol Cell* **7**, 273-82 (2001).
60. Zhao, W. et al. Promotion of BRCA2-Dependent Homologous Recombination by DSS1 via RPA Targeting and DNA Mimicry. *Mol Cell* **59**, 176-87 (2015).
61. Calhoun, S.F. et al. Chromosome End Repair and Genome Stability in Plasmodium falciparum. *MBio* **8** (2017).
62. Lopez-Rubio, J.J., Mancio-Silva, L. & Scherf, A. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe* **5**, 179-90 (2009).
63. Flueck, C. et al. Plasmodium falciparum heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. *PLoS Pathog* **5**, e1000569 (2009).
64. Mattei, D. & Scherf, A. Subtelomeric chromosome instability in Plasmodium falciparum: short telomere-like sequence motifs found frequently at healed chromosome breakpoints. *Mutat Res* **324**, 115-20 (1994).
65. Scherf, A. & Mattei, D. Cloning and characterization of chromosome breakpoints of Plasmodium falciparum: breakage and new telomere formation occurs frequently and randomly in subtelomeric genes. *Nucleic Acids Res* **20**, 1491-6 (1992).
66. Putnam, C.D., Pennaneach, V. & Kolodner, R.D. Chromosome healing through terminal deletions generated by de novo telomere additions in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **101**, 13262-7 (2004).
67. Gupta, D.K., Patra, A.T., Zhu, L., Gupta, A.P. & Bozdech, Z. DNA damage regulation and its role in drug-related phenotypes in the malaria parasites. *Sci Rep* **6**, 23603 (2016).